

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-129470

(43)Date of publication of application : 19.05.1995

(51)Int.Cl.

G06F 12/08

G06F 3/06

(21)Application number : 05-279249

(71)Applicant : HITACHI LTD

(22)Date of filing : 09.11.1993

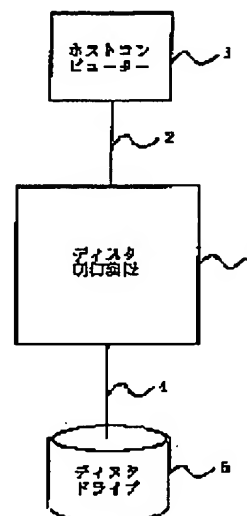
(72)Inventor : FUJII TETSUHIKO  
YAMAMOTO AKIRA  
KOBASHI TETSUZO  
ACHIWA KIYOUSUKE

## (54) DISK CONTROL METHOD

## (57)Abstract:

PURPOSE: To prevent the degradation of sequential read performance and a look-ahead effect by determining the storage location on the disk of compressed data or a physical address not so as to impair the sequence of a logical address and the physical address.

CONSTITUTION: In this method, a host computer 1, a disk controller 3, a disk drive 5, and buses 2 and 4 are provided. The disk controller 3 selects and compresses the data in the dirty extension state of a cache memory and the latest access time becomes old or LRU data and next determines the storage device on the disk of compressed data or a physical address not so as to impair the sequence of a logical address and the physical address as much as possible by using the correspondence of the logical address and the physical address and control information controlling the free area on the disk. The storage location on the disk of the data compressed in this way is determined not so as to impair the sequence of the logical address and the physical address.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the  
examiner's decision of rejection or application  
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of  
rejection][Date of requesting appeal against examiner's decision  
of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-129470

(43) 公開日 平成7年(1995)5月19日

(51) Int.Cl. <sup>9</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 12/08	3 2 0	7608-5B		
3/06	3 0 1 W			

審査請求 未請求 請求項の数16 O L (全 30 頁)

(21) 出願番号 特願平5-279249

(22) 出願日 平成5年(1993)11月9日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 藤井 哲彦

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(72) 発明者 山本 彰

神奈川県川崎市麻生区王禅寺1099番地 株

式会社日立製作所システム開発研究所内

(72) 発明者 小橋 徹三

神奈川県小田原市国府津2880番地 株式会

社日立製作所ストレージシステム事業部内

(74) 代理人 弁理士 小川 勝男

最終頁に続く

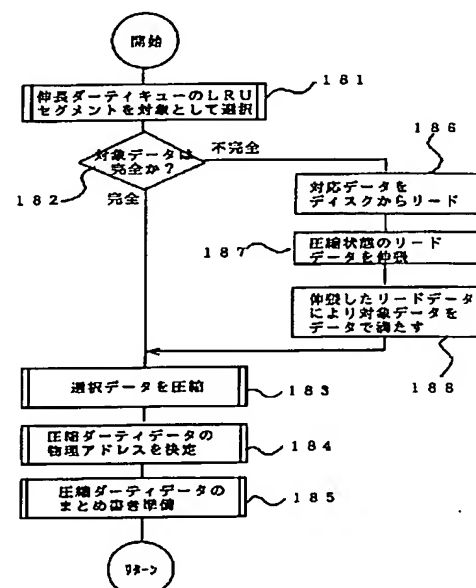
(54) 【発明の名称】 ディスク制御方法

(57) 【要約】

【構成】 キャッシュメモリ上のダーティデータを圧縮した後、データのホストコンピュータに見える論理アドレスと、ディスク上の格納位置である物理アドレスとの、順序性ができるだけ保たれるようなアルゴリズムにより、圧縮したデータの物理アドレスを決定する（ステップ184）。またその後、圧縮したダーティデータを、その物理アドレスに基づき、まとめ書き用のキューにつなぎかえる（ステップ185）。

【効果】 実装するディスクドライブの物理的な記憶容量以上の記憶容量を、シーケンシャルリード性能と先読み効果の低下を防止し、まとめ書きを可能としつつ、提供することが可能となる。

図 18



## 【特許請求の範囲】

【請求項1】 ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、キャッシュメモリ中のダーティな伸長状態の、最新アクセス時刻が古くなった、即ちLRUデータを選択し、圧縮するステップと、論理アドレスと物理アドレスの対応と、ディスク上の空き領域を管理する管理情報を用いて、前記圧縮するステップで圧縮したデータのディスク上の格納位置、即ち、物理アドレスを、論理アドレスと物理アドレスの順序性ができるだけ損なわれないように、決定するステップとを含むことを特徴とするディスク制御方法。

【請求項2】 請求項1において、ディスク上の物理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、前記圧縮されたダーティデータが存在する場合、当該データの管理情報へのポインタを管理するテーブルを用いて、圧縮したダーティデータの物理アドレスに従って、前記テーブルをセットするステップとを含むディスク制御方法。

【請求項3】 請求項1において、圧縮データのディスク上の格納位置を、ディスク上の前記圧縮データに対応する更新前の圧縮データの位置を基点として、圧縮データのデータ長が更新前の圧縮データのデータ長以下である時には、更新前の圧縮データの格納位置と同一とし、圧縮データのデータ長が、更新前の圧縮データのデータ長よりも大きい時には、更新前の圧縮データのディスク上直後の空き領域の有無と、空き領域の長さのチェックを行い、後方の空き領域を更新前の圧縮データに接続して一つの連続領域とした場合に、圧縮データが、当該領域内に収まる場合には、圧縮データのディスク上の格納位置を、更新前の圧縮データの格納位置と同一とし、圧縮データが、領域内に収まらない場合には、更新前の圧縮データのディスク上の格納位置の近傍から、圧縮データを格納可能な大きさの空き領域を捜して、その空き領域を、圧縮データのディスク上の格納位置とするディスク制御方法。

【請求項4】 請求項3において、キャッシュ上のダーティデータを圧縮し、その圧縮データ長が、ディスク上の対応する旧圧縮データのデータ長よりも短くなった時に、新しい圧縮データを、旧圧縮データ上に格納するが、その際に、旧圧縮データが格納されていたディスク上の領域のうち、新しい圧縮データが格納されなかった、後部の領域の全てまたは一部を、空き領域とするディスク制御方法。

【請求項5】 請求項3において、キャッシュ上のダーティデータを圧縮し、その圧縮データ長が、ディスク上の対応する旧圧縮データのデータ長よりも短くなった時に、新しい圧縮データを、旧圧縮データ上に格納するが、その際に、旧圧縮データが格納されていたディスク上の領域のうち、新しい圧縮データが格納されなかった、後部の領域の全てまたは一部を、将来の圧縮データ

長の増加に備えて、空き領域としないで、確保したままにするディスク制御方法。

【請求項6】 請求項1において、キャッシュ上のダーティデータを圧縮し、その圧縮データ長が、ディスク上の対応する旧圧縮データのデータ長よりも長くなった時に、旧圧縮データの直後に、旧圧縮データと接続しても、新しい圧縮データを格納するに十分な大きさの空き領域がないときに、旧圧縮データの後方のデータを、ディスク上の別の位置にある領域に移動することによって、新しい圧縮データを、旧圧縮データの格納位置に格納可能な大きさの、空き領域を作成し、新しい圧縮データを、旧圧縮データの格納位置に格納するディスク制御方法。

【請求項7】 請求項1において、キャッシュ上のダーティデータを圧縮し、その圧縮データ長が、ディスク上の対応する旧圧縮データのデータ長よりも長くなった時に、旧圧縮データの直後に、旧圧縮データと接続しても、新しい圧縮データを格納するに十分な大きさの空き領域がないときに、空き領域を集めるコンパクション処理を起動するディスク制御方法。

【請求項8】 請求項1において、圧縮したキャッシュ上のダーティデータを、ディスクに格納する時に、キャッシュ上のダーティデータを分割して、ディスク上に分散して格納することを許すディスク制御方法。

【請求項9】 請求項1において、ディスクの論理アドレスから物理アドレスを引くことができるテーブルを利用することによって、論理アドレスと物理アドレスの対応を管理するディスク制御方法。

【請求項10】 請求項1において、ディスク上の空き領域を、ディスクのブロック毎に、前記ブロックが空き状態か否かをビット値によって示すビットマップ情報を用いて管理するディスク制御方法。

【請求項11】 請求項1において、ディスク上の空き領域を、空き領域のサイズ毎に管理することにより、必要な大きさの空き領域の検索を高速に行うディスク制御方法。

【請求項12】 請求項1において、圧縮データのディスク上の格納位置を、データの論理的なアドレスに、ディスク内のデータの平均圧縮率を乗じた値を、圧縮データのディスク上の格納位置の基点アドレスとし、基点アドレスの近傍から、空き領域を検索して、圧縮データのディスク上の格納位置とするディスク制御方法。

【請求項13】 請求項1において、圧縮されたキャッシュ上のダーティデータを、ディスクにまとめ書きしようとする時に、ディスク上の論理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、圧縮されたダーティデータが存在する場合、そのディスクへの格納位置即ち物理アドレスとを管理するテーブルを用いて、まず論理アドレス上近傍の圧縮されたダーティデータを選び出すステップと、前記選び出すステップで選び出した圧

縮されたダーティデータのうち、前記テーブルにより、物理アドレス上近傍のものだけを選択するステップにより、まとめ書き対象の、圧縮されたダーティデータを選択することにより、キャッシュ上のディスク未反映のダーティデータのまとめ書きを可能とするディスク制御方法。

【請求項 14】請求項 2 において、圧縮されたダーティデータをディスクへライトしようとする時に、圧縮されたダーティデータのディスクへの格納位置を管理するテーブルを用いて、次にディスクライトすべき圧縮されたダーティデータを定めるディスク制御方法。

【請求項 15】請求項 2 において、圧縮されたダーティデータをディスクへライトしようとする時に、圧縮されたダーティデータのディスクへの格納位置を管理するテーブルと、現在のヘッドのシリンダ位置情報とを併用して、次にディスクライトすべき圧縮されたダーティデータを定めるディスク制御方法。

【請求項 16】請求項 2 において、圧縮されたダーティデータをディスクへライトしようとする時に、圧縮されたダーティデータの圧縮した時刻順、あるいは圧縮前の伸長状態のダーティデータへのホストコンピュータからの最終アクセス時刻が古い順、あるいは圧縮前の伸長状態のダーティデータを管理するキュー中のデータの順序を用いることによって、次にディスクライトすべき圧縮されたダーティデータを定めるディスク制御方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明はデータ圧縮機能を有するキャッシュ付きディスク制御装置のディスク制御方法に関する。

【0002】

【従来の技術】第一の文献シージーエスグループ：ザエバリエーター シリーズ、アン アナリシス オブザ ストレージテック アイスバーク 9200 ストレージシステム (CGS GROUP: The Evaluator Series, An Analysis Of The StorageTekIceberg 9200 Storage System, ES-0003-0, March 1992, 2-9.) は、データ圧縮機能とライトバックキャッシュを有する、ディスクアレイサブシステムについて述べている。このディスクアレイサブシステムでは、ホストインタフェースとキャッシュメモリとの間にデータ圧縮部を持ち、ホストコンピュータから送られてくるライトデータを、圧縮し、圧縮したデータをキャッシュメモリに格納する。ライトデータのディスクへの格納は、ライトアクセスの順序に、データを追記形式で格納する。

【0003】特開平3-172946 号公報では、ライトバックキャッシュ制御を行うディスク制御装置で、キャッシュメモリ上のダーティデータを、ディスクにライトするときに、ディスク上で互いに近い位置にあるダーティデータを、一連のディスクライト処理でまとめてディスク

ライトすることにより、ライトバック動作に要するシーク、サーチといったディスクのメカニカルな動作時間を節約して、性能向上を図る方法として、ディスクライトしようとする一つのダーティデータ（ダーティデータ a とする）と同一のシリンダ内にあるダーティデータを、各シリンダのダーティデータの有無等を管理するシリンダ管理情報と、キャッシュがアロケートされているシリンダに対して、シリンダ内の各トラックのダーティデータ有無等を管理するトラック管理情報とにより、把握することにより、ダーティデータ a とまとめ書きを行うダーティデータを求め、短時間で、多くのダーティデータをディスクライトする方法を開示している。

【0004】第二の文献アンドリュウ・ダブリュー・アペル アンド カイ・リー：バーチャル メモリ プリミティブズ フォー ユーザー プログラムズ、アーキテクチャ サポート フォー プログラミング ランゲージ アンド オペレーティング システムズ、ピーピー 96-107, サンタクララ, カリフォルニア, エプリール 8-11, 1991 (Andrew W. Appel and Kai Li: Virtual Memory Primitives for User Programs, Architectural Support for Programming Language and Operating Systems, pp. 96-107, Santa Clara, California, April 8-11, 1991) は、コンピュータの仮想記憶システムにおいて、キャッシュメモリとデータ圧縮を組合せ、キャッシュメモリに伸長データと圧縮データを混在させるという技術思想を示している。それは次のようなものである。仮想記憶システムで、しばらくの間アクセスされていないページ (less-recently-used pages) は、ディスクにページアウトするかわりに、データを圧縮して主記憶に格納しておく。圧縮して主記憶に格納されたページが再び必要になったときには、それを伸長すれば良く、ディスクから取り出すのに比べて短時間で済む。更に、より長期間にわたって使用されなかった、圧縮されたページは、ディスクに送出する。

【0005】

【発明が解決しようとする課題】ディスク制御装置において、第一の文献に示されているように、ライトバックキャッシュとデータ圧縮機能を、組合せて採用することにより、ディスクキャッシュによる応答時間の高速化と、データ圧縮機能によるディスクの実効記憶容量の向上を図ることができる。ここで、ディスクの実効記憶容量を向上させるために、ディスクには圧縮状態のデータを記憶する。この場合、圧縮アルゴリズムの一般的な性質として、データの内容によって圧縮されたデータの長さが変動する。あるデータ（データ 1 とする）がディスク上で圧縮されて格納されており、その格納位置の直後に別のデータ（データ 2 とする）が圧縮されて格納されていたとする。ここでデータ 1 は圧縮されて、長さが 0 バイトであったとする。ここでデータ 1 が書き換えら

れてデータ1'になったとする。データ1'をディスクに記憶するために圧縮したデータの長さが、 $L1$ バイト ( $L1 > L0$ ) になったとする。この場合、データ1の圧縮データが格納されていた領域は、ちょうど $L0$ バイトであったので、データ1'を圧縮したデータは、もとの位置に収まりきらない。そこで、データ1'をディスク上の別な位置にある $L1$ バイト以上の空き領域に格納するか、データ2を別の位置に移動しなければならない。すなわち、データ圧縮機能により、実効記憶容量を向上させようとするディスク装置では、データが更新されると、ディスク上の対応する圧縮データの格納位置を、動的に変更しなければならないという事態が生じる。

【0006】第一の文献に示される方法では、ホストインタフェースとキャッシュメモリとの間にデータ圧縮部を持ち、ホストコンピュータから送られてくるライトデータを、圧縮し、圧縮したデータをキャッシュメモリに格納し、ライトデータのディスクへの格納は、ライトアクセスの順序に、データをディスク上の空き領域へ追記形式で格納することにより、前記の事態に対処している。

【0007】しかし、第一の文献の方法では、圧縮データを、ライト順にディスクに追記するため、データの論理アドレスと対応する圧縮データのディスク上の格納位置との順序性が損なわれ、シーケンシャルリード性能や先読み効果が低下するという課題があった。

【0008】また、第二の文献では、仮想記憶に通常の伸長状態のページと、圧縮されたページを混在させる方法を述べているが、圧縮されたページを、どのような方法で、ディスクに格納するかについては、詳細な記述がされていない。

【0009】(1) 本発明の第一の課題は、ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、ホストコンピュータからみえる論理的なデータのアドレスと、対応する圧縮データのディスク上の格納アドレスとの間の順序性をできるだけ保つことにより、シーケンシャルリード性能や先読み効果の低下を防止することである。

【0010】更に、ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、ホストコンピュータからみえる論理的なデータのアドレスと、対応する圧縮データのディスク上の格納アドレスとの間の順序性をできるだけ保つことにより、シーケンシャルリード性能や先読み効果の低下の防止を図った場合において、キャッシュメモリ上のダーティデータのディスクへのまとめ書き機能を実現しようとした場合に、特開平3-172946号公報の方法では、キャッシュメモリ上のダーティデータのディスク上の格納位置が動的に変化することが考慮されておらず、そのままでは、ダーティデータのまとめ書きを行うことができないという課題があっ

た。また、第二の文献には、圧縮データのディスクへの格納方法について、詳細な記述がされていない。

【0011】(2) 本発明の第二の課題は、ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、第一の課題を解決すると共に、更に、キャッシュ上のディスク未反映のダーティデータのまとめ書きを可能とすることである。

【0012】

【課題を解決するための手段】第一の課題を解決するために、本発明はキャッシュメモリ中のダーティな伸長状態の、最新アクセス時刻が古くなった、即ち、LRUデータを選択し、圧縮するステップ(ステップ1)と、論理アドレスと物理アドレスの対応と、ディスク上の空き領域を管理する管理情報を用いて、前記ステップ1で圧縮したデータのディスク上の格納位置即ち物理アドレスを、論理アドレスと物理アドレスの順序性ができるだけ損なわれないように、決定するステップ(ステップ2)とを設ける。

【0013】第二の課題を解決するために、本発明は第一の課題を解決するためのステップ1、2に加えて、ディスク上の物理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、圧縮されたダーティデータが存在する場合、データの管理情報へのポインタとを管理するテーブルを用いて、ステップ1、2に続いて、圧縮したダーティデータの物理アドレスに従って、前記テーブルをセットするステップ(ステップ3)とを設ける。

【0014】また、第二の課題を解決するために、本発明は第一の課題を解決するためのステップ1、2に加えて、ディスク上の論理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、圧縮されたダーティデータが存在する場合、そのディスクへの格納位置、即ち、物理アドレスとを管理するテーブルを用いて、まず論理アドレス上近傍の圧縮されたダーティデータを選び出すステップ(ステップ4)と、更に前記ステップ4で選び出した圧縮されたダーティデータのうち、前記テーブルにより、物理アドレス上近傍のものだけを選択するステップ(ステップ5)とを設ける。

【0015】

【作用】ディスク制御装置は、ステップ1によりキャッシュメモリ中のダーティな伸長状態の、最新アクセス時刻が古くなった、即ち、LRUデータを選択して圧縮し、次にステップ2により論理アドレスと物理アドレスの対応と、ディスク上の空き領域を管理する管理情報を用いて、ステップ1で圧縮したデータのディスク上の格納位置即ち物理アドレスを、論理アドレスと物理アドレスの順序性ができるだけ損なわれないように決定する。ステップ2で、圧縮したデータのディスク上の格納位置を、論理アドレスと物理アドレスの順序性ができるだけ損なわれないように決定するので、シーケンシャルリード性能や先読み効果の低下を防止することができる。

【0016】ディスク制御装置は、ステップ1、2に引き続き、ステップ3によりディスク上の物理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、圧縮されたダーティデータが存在する場合、当該データの管理情報へのポインタとを管理するテーブルを用いて、圧縮したダーティデータの物理アドレスに従って、テーブルをセットする。

【0017】ディスク制御装置は、ステップ3でセットされるテーブルを用いて、圧縮されたダーティデータの、ディスクへのまとめ書きを行う。このまとめ書きのステップは、シリンダのスケジューリングを、例えば、いわゆるSSTF (Shortest Seek Time Fast) アルゴリズムや、SCANアルゴリズムを使用し、同一シリンダ内のダーティデータのスケジューリングを、例えば特開平3-172946号公報に記載の方法によって実現することができる。これにより、第一の課題を解決すると共に、更にキャッシュ上のダーティデータのディスクへのまとめ書きを可能とすることができる。

【0018】ディスク制御装置は、ステップ1、2の後、ステップ4によりディスク上の論理アドレス毎に、圧縮されたダーティデータの有無を表す情報と、圧縮されたダーティデータが存在する場合、そのディスクへの格納位置即ち物理アドレスとを管理するテーブルを用いて、まず論理アドレス上近傍の圧縮されたダーティデータを選び出し、更にステップ5により前記ステップ4で選び出した圧縮されたダーティデータのうち、テーブルにより、物理アドレス上近傍のものだけを選択する。それからディスク制御装置は、選択した、物理アドレス上近傍の、キャッシュ上のダーティな圧縮データを、ディスクへまとめ書きする。これにより、第一の課題を解決すると共に、更にキャッシュ上のダーティデータのディスクへのまとめ書きを可能とすることができる。

【0019】

【実施例】本発明を実施した、圧縮機能とライトバックキャッシュを有する、ディスク制御装置について説明する。本実施例では、古くなったダーティデータを、ディスクライトに備えて圧縮し、ディスク上の格納位置（物理アドレス）を決めた後、その物理アドレスに従って、物理アドレスから圧縮したダーティデータを、検索可能なテーブル下のキューにつなぎ、まとめ書き対象を、そのテーブルを検索することによって求める。

【0020】最初に、基本的な方式を述べる。

【0021】本実施例では、ディスクコントローラとディスク間のデータ転送は圧縮データで行い、ディスクコントローラとホストコンピュータの間のデータ転送は伸長データで行うことを前提とする。即ち、ホストコンピュータは、伸長データを扱う。また、ディスクには圧縮データを格納する。キャッシュメモリには、伸長データと圧縮データを混在させる。ホストコンピュータから最

近アクセスされた、再アクセス性の高いデータは、リードヒット時の高速応答のために、伸長状態でキャッシュに置き、しばらくの間アクセスされていない (less recently used) ダーティデータは、ディスクへのライトに備えて、キャッシュ上で圧縮し、圧縮状態でキャッシュに置く。しばらくの間アクセスされていないクリーンデータについては、キャッシュ上に伸長状態で置く方法と、圧縮状態で置く方法とがあるが、ここではキャッシュ上に伸長状態で置くものとする。

10 【0022】次に、使用する用語の定義を行う。

【0023】・論理ディスク：ホストコンピュータから見えるディスクを論理ディスクと呼ぶ。圧縮機能があるので、実際のディスクの容量より大きな容量を持つ。

15 【0024】・物理ディスク：実際にデータを格納するディスクを物理ディスクと呼ぶ。

【0025】・論理ディスクセグメント：論理ディスクの管理単位。

【0026】・コンプレッド論理ディスクセグメント：圧縮された論理ディスクセグメント。

20 ・伸長データ：圧縮する前の、通常のデータのこと。

【0027】・圧縮データ：圧縮されたデータのこと。

25 【0028】・ダーティ/クリーン：ディスクに未反映のライトデータがキャッシュ上にある状態をダーティと呼ぶ。クリーンとは、データがすべてディスクに反映済みのこと。

【0029】・キャッシュセグメント：キャッシュメモリの管理単位。

30 【0030】・物理ディスクシリンダグループ：まとめ書きのための物理ディスクの管理単位であり、ここでは、物理ディスク上の1～n個の連続したシリンダを物理ディスクシリンダグループとする。

【0031】・物理ディスクセグメント：物理ディスクの管理単位。p (≧1) セクタとする。本実施例では、1セクタを物理ディスクセグメントとする。

35 【0032】・完全/不完全：セグメント内がすべて有効データで満たされている場合を完全と呼び、そうでない場合を不完全と呼ぶ。不完全なデータはこのままでは圧縮できない。

40 【0033】図1は、本発明を適用するディスク制御装置の位置付けを示す。図1中、1がホストコンピュータ、3がディスク制御装置、5がディスクドライブ、2はホストコンピュータとディスク制御装置を結ぶバス、4はディスク制御装置とディスクドライブを結ぶバスである。

45 【0034】図2は、本実施例におけるディスク制御装置のハードウェア構成を示す。図2中、301はディスク制御装置内部のバス、302はホストインタフェース、304はMPU (メインプロセッシングユニット)、306はMPUの主メモリ、308はDMAC (ダイレクトメモリアccessコントローラ)、313は

データの圧縮伸長を行う圧縮復元部、315はHDC（ハードディスクコントローラ）、32はDMACが制御するメモリ、33はメモリ内のキャッシュ/バッファ部分、303、305、307、309、310、314、316は、それぞれデータやコマンド等を受受する信号線である。ディスク制御プログラムは、主メモリ306に置かれ、MPU304が実行する。キャッシュメモリの管理情報等の制御情報は、主メモリ306および/またはメモリ32に置かれる。DMAC308は、メモリ32へのアクセス制御を行う。HDC315はディスクドライブ5の制御を行う。圧縮復元部313は、データの圧縮及び伸長を行う。

【0035】図5は、論理ディスク管理テーブルである。ここで論理ディスクとは、ホストコンピュータに見えるディスクを指す。論理ディスクの記憶容量は、実際のディスクの記憶容量より大きくなる。図5中、79は一つの論理ディスクセグメントに対応するテーブルエントリである。Ndsは、論理ディスクセグメントの総個数である。71はヒットミス判定情報、72はキャッシュセグメント番号、73は当該論理ディスクセグメントが使用中か未使用かを示すフラグ、74は当該論理ディスクセグメントのデータが、物理ディスク上圧縮されているか否かを示すフラグ、75と76はそれぞれ、当該論理ディスクセグメントに対応する圧縮データの、物理ディスク上の先頭アドレスとレンジ、77と78はそれぞれ、当該論理ディスクセグメントに対応する圧縮データの、物理ディスク上の、最も近い前方、及び後方の圧縮された論理ディスクセグメントへのポインタである。

【0036】図6は、キャッシュ管理テーブルである。図6中、80は一つのキャッシュセグメントに対応するテーブルエントリである。Ncsは、キャッシュセグメントの総数である。81はキャッシュセグメントが使用中か否かを示すフラグ、82はキャッシュセグメント中のデータが圧縮データか伸長データかを示すフラグ、83はキャッシュセグメントがダーティか否かを示すフラグ、84はキャッシュセグメント中の各ブロックに有効データが存在しているか否かを示すビットマップ情報、85はキャッシュセグメント中の各ブロックにダーティデータが存在しているか否かを示すビットマップ情報、86はキャッシュセグメント中のデータが圧縮データであった場合、そのデータ長を示すデータ、87はキャッシュセグメントへの最新のアクセスのアクセス時刻値、88と89は、それぞれキャッシュセグメントが存在するキュー中での、直前及び直後のキャッシュセグメントをポイントする前方ポインタと後方ポインタ、8Aは、キャッシュセグメントに割当てられている論理ディスクセグメントの番号である。

【0037】図11は、メモリ32の構成である。図11中、34は制御テーブル、330はキャッシュメモリ

部、331はディスクリードデータ格納バッファ、332は伸長用ワークバッファ、333は確認用ワークバッファである。

【0038】図7は、キャッシュのデータ格納部を示す。図中91は、一つのキャッシュセグメントである。

【0039】図8は、キャッシュキュー管理テーブルである。図8中、93は空き状態のキャッシュセグメントのリストの先頭ポインタ、94と95はそれぞれ、伸長状態かつクリーン状態の、キャッシュセグメントのLRUキューのMRUポインタとLRUポインタ、96と97はそれぞれ、伸長状態、かつ、ダーティ状態の、キャッシュセグメントのLRUキューのMRUポインタとLRUポインタ、98は、圧縮ダーティデータのまとめ書き用のテーブルであり、物理ディスクを1～nシリンダの物理ディスクシリンダグループに分割し、それぞれの領域毎の、その領域に格納されるべき物理アドレスを持つ圧縮ダーティキャッシュセグメントの、物理アドレス順のキューの先頭ポインタである。

【0040】図9は物理ディスク内の空き領域を管理するテーブルである。図9のテーブルを用いた場合、空き領域は、領域長毎に管理する。図9中、101は領域長が(i-1)ディスクセグメント以上かつiディスクセグメント未満の、連続空き領域のキューの先頭ポインタである。

【0041】図10は、空き領域管理用のエントリであり、103と104はそれぞれ、空き領域の先頭アドレスと領域長である。

【0042】次に、図12ないし図23、図4のフローチャートを用いて、ディスク制御装置の基本的な動作を説明する。

【0043】図12は、ディスクの初期化処理のフローである。ディスクの初期化処理は、ホストコンピュータ1からの初期化コマンドによって起動される。ディスク制御装置3のMPU304は、ホストコンピュータ1からの初期化コマンドを受け取ると、初期化処理を開始する。初期化処理の一例を次に示す。まず、ステップ121でキャッシュメモリ330を0クリアし、次にステップ122で物理ディスク（＝ディスクドライブ5）を初期化し、次にステップ123で論理ディスク管理テーブル7を初期化し、次にステップ124でキャッシュ管理テーブル8を初期化し、次にステップ125でキャッシュキュー管理テーブル92を初期化し、次にステップ126で空き領域管理テーブル100を初期化し、次にステップ127でバッファ331、332、333を初期化し、次にステップ128で全ての論理ディスクセグメントに対して初期設定をする。

【0044】ステップ123の処理では、各論理ディスクセグメントのエントリ79に対して、chm71をミスに、csn72をnull値に、ldu73を未使用に、ldst74、hd\_\_addr75、dlen76、d



`_fwd_ptr77`, `d_bwd_ptr78`の値をすべてnull値にする。

【0045】ステップ124の処理では、各キャッシュセグメントのエントリ80に対して、`csu`を未使用に、`csst`をnull値に、`csd`をノットダーティに、`exist_bits`をオールオフに、`dirty_bits`をオールオフに、`clen`をnull値に、`access_time`に $-\infty$ を、`c_fwd_ptr`と`c_bwd_ptr`をそれぞれ、キャッシュセグメント*i*に対して、キャッシュセグメント(*i*-1)とキャッシュセグメント(*i*+1)のキャッシュセグメント番号をセットする。なお、キャッシュセグメント0の`c_bwd_ptr`と、キャッシュセグメント*Ncs*の`c_fwd_ptr`はそれぞれnull値をセットする。

【0046】ステップ125の処理では、図8のキャッシュキュー管理テーブルを初期化するが、初期状態では、キャッシュセグメントはすべて空き状態である。したがって、伸長クリーンキャッシュセグメントのLRUキューのMRUポインタ94及びLRUポインタ95、伸長ダーティキャッシュセグメントのLRUキューのMRUポインタ96及びLRUポインタ97はnull値とする。また、物理ディスクシリンダグループ毎の、圧縮ダーティキャッシュセグメントキューの先頭ポインタ98もnull値とする。空きキャッシュセグメントリストの先頭ポインタ93は、キャッシュセグメント0のセグメント番号を代入する。

【0047】ステップ126の処理では、空き領域の管理テーブルを初期化する。ここでは、すべての物理ディスク上の領域が空き状態であるものとしてテーブル初期化を行う。

【0048】ステップ128の処理では、全論理ディスクセグメントの初期設定を行う。例えばオールゼロや、オールブランクのような初期化データで、論理ディスクのセクタを初期化する。この方法は、以下に述べる、ホストコンピュータ1からのライトコマンドの処理と同様に行うことができる。また、実際に初期化データを、物理ディスクにライトせずに、ディスク制御装置3で、初期化データのエミュレーションを行うことも可能である。この場合、論理ディスクの各セクタへの、初期化以後の、ライトの有無をビット等で管理し、リードに対して、ライト有りであれば、通常のリードとして処理し、ライト無しであれば、ディスク制御装置3が、初期化データを、ホストコンピュータ1へ返せば良い。

【0049】次に、ホストコンピュータ1からディスク5へのリード/ライトコマンドの処理について述べる。図13は、リード/ライトコマンド処理の全体フローである。まず、コマンドがリードかライトかを判定する(ステップ131)。ライトであれば、ステップ132へ進み、ヒットミス判定を行う。ヒットミス判定は、ホストコンピュータ1が指定するデータのアドレスから、

データの属する論理ディスクセグメント番号を求め、論理ディスクセグメントの論理ディスク管理テーブル7の対応エントリ79の、ヒットミス判定情報`chm71`をチェックすることにより可能である。ヒットであれば、ライトヒット処理(ステップ133)を行い、ミスであれば、ライトミス処理(ステップ135)を行い、処理を終了する(ステップ134)。ステップ131で、コマンドがリードであれば、ステップ136へ進み、ヒットミス判定を行う。ヒットであれば、リードヒット処理(ステップ137)を行い、ミスであれば、リードミス処理(ステップ138)を行い、処理を終了する(ステップ134)。

【0050】以下、ライトヒット/ミス、リードヒット/ミスの各処理を詳細に説明する。ライトミス処理のフローを図14に示す。まず、ミスした論理ディスクのデータに対して、キャッシュセグメントを新たに割当てて(ステップ141)、ホストコンピュータからの伸長状態のライトデータをそのまま、割当てたキャッシュセグメントへライトする(ステップ142)。それから、キャッシュセグメントを、伸長ダーティキャッシュセグメントキューのMRU位置につなぎ(ステップ143)、IO完了をホストコンピュータへ報告し(ステップ144)、リターンする。キャッシュセグメントには、伸長状態のライトデータが格納されているので、キャッシュセグメントは、伸長ダーティ状態となる。ステップ141のキャッシュセグメントの割当ては、キャッシュセグメントの空きキューを、先頭ポインタ93から順次検索し、空きがなければ、キャッシュセグメントのクリーンキューを、LRUポインタ95から順次検索することによって、リプレースするキャッシュセグメントを選択すれば良い。

【0051】ライトヒット処理のフローを図15に示す。まず、ヒットデータが伸長状態か、圧縮状態かを判定する(ステップ151)。ヒットデータが圧縮状態であれば、圧縮状態のヒットデータを伸長して、キャッシュセグメントを圧縮し(ステップ155)、ステップ152へ進む。これは、キャッシュメモリ33から、DMAC308、バス301を介し、圧縮復元部313に、圧縮データを送り、圧縮復元部313で、データを復元し、バス301、DMAC308を介して、伸長データを、キャッシュメモリ33に返すことにより可能である。ヒットデータが伸長状態であれば、そのままステップ152へ進む。ステップ152で、ライトデータを、伸長状態のままキャッシュメモリへ書き込み、ステップ153へ進む。ステップ153では、キャッシュセグメントキューの更新を行う。圧縮状態のデータにヒットしていた場合には、キャッシュセグメントを、圧縮ダーティキューからはずして、伸長ダーティキューのMRU位置につなぐ。伸長状態のデータにヒットしていた場合には、キャッシュセグメントを、それが属していたキュー



一、即ち、伸長クリーンキューもしくは、伸長ダーティキューからはずして、伸長ダーティキューのMRU位置につなぐ。それからステップ154へ進み、ホストコンピュータ1にI/O完了を報告し、リターンする。ここではライトヒットデータが圧縮状態であった場合に、伸長してから、データをライトするようにしているが、ライト範囲が圧縮伸長単位であった場合には、伸長せずに、そのまま書きすることも可能である。

【0052】リードヒット処理のフローを図16に示す。まず、ヒットデータが伸長状態か圧縮状態かを判定する(ステップ161)。ヒットデータが圧縮状態であった場合、圧縮状態のリードヒットデータを伸長し(ステップ165)、ステップ162へ進む。ヒットデータが伸長状態であれば、そのままステップ162へ進む。ステップ162で、データをキャッシュメモリ33からホストコンピュータ1へ転送し、ステップ163へ進む。ステップ163ではキューの更新を行う。ステップ163では、キャッシュセグメントを、それが属していたキューからはずし、ヒットデータがクリーン状態であれば、伸長クリーンキューのMRU位置へ、ヒットデータがダーティ状態であれば、伸長ダーティキューのMRU位置へつなぐ。それから、ステップ164へ進み、I/O完了をホストコンピュータ1に報告し、リターンする。

【0053】リードミス処理のフローを図17に示す。まず、キャッシュセグメントを新たに割当て(ステップ171)。割当てはライトミスと同様に行えばよい。この場合、リードミスであるから、リードデータはディスクに格納されている。しかし、ディスク上の格納位置と、ホストコンピュータ1が指定してくるリードデータのアドレスは、一般的には一致しない。これは、ディスク制御装置3で、データ圧縮を行っており、ホストコンピュータ1の指定するアドレスと、実際のディスク上の格納位置との対応付けが動的に変化するためである。そこで、ディスク上の格納位置(アドレス)を、ホストコンピュータ1が指定するリードデータのアドレスを基に、論理ディスク管理テーブル7から算出する。具体的には、ホストコンピュータ1が指定するリードデータのアドレスから、論理ディスクセグメント番号を求め、論理ディスクセグメントのエントリのhd\_addr75をテーブル7から求めることによって、求めることができる(ステップ172)。次に、ステップ172で求めた、ディスク上の格納位置を基に、リードデータに対応する圧縮データをディスク5から、キャッシュメモリ33へリードする(ステップ173)。次に、リードした圧縮データを、キャッシュメモリ33上で伸長し、ステップ171で確保したキャッシュセグメントへデータを格納する(ステップ174)。次に、伸長したデータのうち、リード対象部分を、キャッシュメモリ33からホストコンピュータ1へ転送する(ステップ175)。次

に、キャッシュセグメントを、伸長クリーンキャッシュセグメントキューのMRU位置につなぐ(ステップ176)。最後に、ホストコンピュータ1にI/O完了を報告し(ステップ177)、リターンする。

05 【0054】以上、通常のリード/ライト処理について述べた。

【0055】ライトバックキャッシュを有する、圧縮機能付きのディスクでも、通常のライトバックキャッシュのときと同様に、キャッシュメモリ上のディスク未反映のデータ、即ち、ダーティデータのディスクライト処理が必要になる。もし、ダーティデータのディスクライト処理を行わないと、キャッシュメモリがダーティデータでいっぱいになり、その時にキャッシュミスが生じると、キャッシュ割当てをするために、ダーティデータのディスクライト処理を行わなければならないからである。

【0056】以降、ダーティデータのディスクライト処理、即ちライトバック処理について説明する。

【0057】まず、ライトバック処理の概要を説明する。キャッシュメモリ上では、アクセス時刻の新しいデータは、伸長状態で存在する。伸長状態でキャッシュメモリに置かれている、ダーティデータのうち、古くなったものを、選択して、キャッシュメモリ上で圧縮する。圧縮すると、圧縮データの長さが決まるので、圧縮データ長を使用して、ディスク上の格納位置(物理アドレス)を決定する。それから、ディスク上の格納位置に基づいて、まとめ書き用のキュー、すなわち、物理ディスクシリンダグループ毎の圧縮ダーティキャッシュセグメントキューにつなぐ。

30 【0058】以上の処理とは非同期に、物理ディスクシリンダグループ毎の圧縮ダーティキャッシュセグメントキューを基にして、圧縮ダーティデータのディスクへのまとめ書きを行う。以下、図面を用いて、ライトバック処理を詳細に説明する。

35 【0059】ライトバック処理は、ライトバックスケジュール処理と、まとめ書き処理とから構成される。まず、ライトバックスケジュール処理を、図4、図18、図20、図23、図21を用いて説明する。ライトバックスケジュール処理は次のようになる。まず、ステップ201で、キャッシュメモリ33上のダーティデータ量をチェックする。このチェックは、ダーティキャッシュセグメントの個数を、常にカウントしておくことによって可能である。チェックの結果、ダーティデータ量が、閾値 $\alpha h$ 未満であれば、ステップ202へ進む。チェックの結果、ダーティデータ量が、閾値 $\alpha h$ 以上であれば、ステップ204へ進む。

45 【0060】ステップ204では、伸長ダーティデータのディスクライト準備処理を行う。この処理は、後で、図18により詳しく述べる。それから、ステップ205へ進み、再度、キャッシュ上のダーティデータ量のチェ

ックを行う。ダーティデータ量が、閾値  $\alpha 1$  ( $\alpha 1 < \alpha h$  となるように設定する) 以上であれば、ステップ 204 へ戻り、伸長ダーティデータのディスクライト準備処理を継続する。ステップ 205 で、ダーティデータ量が閾値  $\alpha 1$  未満であれば、ステップ 202 へ進む。ステップ 202 では、伸長ダーティキューと、伸長クリーンキューの、それぞれの LRU 位置のキャッシュセグメントの最終アクセス時刻を比較する。それぞれのキューの LRU 位置のキャッシュセグメントは、キャッシュキュー管理テーブル 92 の伸長ダーティキャッシュセグメント LRU キュー LRU ポインタ 97 と、伸長クリーンキャッシュセグメント LRU キュー LRU ポインタ 95 から、セグメント番号を求めることができる。また最終アクセス時刻は、キャッシュ管理テーブル 8 の、access\_time 87 から求めることができる。ステップ 202 の比較の結果、伸長ダーティキューの LRU 位置のセグメントの方が最終アクセス時刻が古ければ、ステップ 206 へ進む。伸長ダーティキューの LRU 位置のセグメントの方が、最終アクセス時刻が新しければ、ステップ 203 へ進む。ステップ 206 で、伸長ダーティデータのディスクライト準備処理を行い、ステップ 202 へ戻る。ステップ 203 では、一定時間待ち、ステップ 201 へ戻る。

【0061】次に、伸長ダーティデータのディスクライト準備の処理について説明する。

【0062】図 18 は、キャッシュメモリ上の、伸長状態のダーティデータのディスクライト準備の処理フローチャートである。まず始めに、伸長ダーティキューの LRU 位置のキャッシュセグメントを、ディスクへライトバックする対象データとして選択する(ステップ 181)。次に、ステップ 181 で選択した伸長ダーティデータが、完全か不完全かを判定する。データが完全であった場合、ステップ 183 へ進む。データが不完全であった場合、ステップ 186 へ進む。ステップ 186 で、選択した対象データに対応する圧縮データのディスク上の格納アドレスを、論理ディスク管理テーブルから求め、ディスクから圧縮データをリードする。次に、ステップ 187 で、リードした圧縮データを伸長する。次に、ステップ 188 で、伸長したリードデータを用いて、ライトバック対象として選択した不完全な伸長ダーティデータを、有効データで満たす。それからステップ 183 へ進む。ステップ 183 では、完全な状態の、選択した伸長ダーティデータを圧縮する。それからステップ 184 で、圧縮したダーティデータのディスク上の格納位置、即ち物理アドレスを決定する。それからステップ 185 で、圧縮したダーティデータをまとめ書き用のキューにつなぎ、まとめ書きの準備をし、リターンする。

【0063】図 20 により、圧縮ダーティデータの物理アドレス決定方法を詳細に説明する。まず、ステップ 21 で、現在ディスクに格納されている、対象データに対

応する圧縮データ(旧圧縮データと呼ぶことにする)のデータ長を、論理ディスク管理テーブル 7 の、dlen 76 から求め、変数 L0 に代入する。一方、ディスクにライトしようとしている対象データを、ステップ 183 で圧縮していた。この、ディスクへのライト対象の圧縮データ(新圧縮データと呼ぶことにする)のデータ長を変数 L1 に代入する。

【0064】それから、ステップ 22 で、L0 と L1 の大きさを比較する。L1 ≤ L0 であった場合、ステップ 23 へ進む。この場合、圧縮したライトデータが、元の圧縮データの大きさ以下であるので、元の領域にそのまま書くことができる。そこで、ステップ 23 では、元の圧縮データ(旧圧縮データ)に、ディスクライトしようとする、圧縮データ(新圧縮データ)を上書きし、それからリターンする。ステップ 22 で、L1 > L0 であった場合、ステップ 24 へ進む。ステップ 24 で、旧圧縮データの直後に、(L1 - L0) 以上の大きさの空き領域があるか否かをチェックする。ここで、論理ディスク管理テーブル 7 から、対象データのディスク上での直後の、圧縮された論理ディスクセグメントの番号を求め、そのディスク上の先頭アドレスを、やはりテーブル 7 から求めることにより、旧圧縮データの直後の空き領域の大きさを求めることができる。(L1 - L0) 以上の大きさの空き領域があった場合、ステップ 25 へ進み、旧圧縮データとその直後の空き領域を、一つの連続した領域とみなして、その先頭から新圧縮データを書き込み、それからリターンする。ステップ 24 で、(L1 - L0) 以上の大きさの空き領域がなかった場合、ステップ 26 へ進み、旧圧縮データの近傍から、L1 以上の大きさの連続した空き領域を検索し、新圧縮データを書き込み、リターンする。前述の空き領域の検索は、例えば、図 3 の、空き領域管理テーブル 600 を使用して、旧圧縮データの位置を基点として、その近傍から、L1 以上の大きさの空き領域をサーチすることによって可能である。また、論理ディスク管理テーブル 7 の、d\_fwd\_ptr、d\_bwd\_ptr を使用して、旧圧縮データの位置を基点として、ディスク上前方および/または後方の、論理ディスクセグメントをポインタで順次たぐってゆき、間にある空き領域の大きさを L1 と比較することによっても可能である。また、図 9 に示した空き領域管理テーブルを用いて、L1 の大きさの連続空き領域を、リストアップし、その中から、元の領域に最も近い、空き領域を見つけて割当て、新圧縮データを書き込むという方法によることも可能である。

【0065】図 23 により、圧縮ダーティデータの物理アドレス決定方法の代替案を説明する。

【0066】まず、ステップ 2301 で、変数 S1 にディスクの物理容量をセットし、変数 S2 に圧縮機能付きディスクの、ホストコンピュータに見せる容量、即ち、公称容量をセットする。ここでは、公称容量は、物理容

量をデータの平均圧縮率（データが1/2になったとき、平均圧縮率は0.5とする。）で割った値とする。さらに変数xに、今ディスクにライトしようとしている、伸長ダーティデータの論理アドレスをセットする。さらに、 $y = x \times S1 / S2$ とする。ここで、yは、論理アドレスxのデータを、物理ディスクのどの位置に格納しようかという、目標の物理アドレスである。次に、ステップ2302で、物理ディスク上で、アドレスyになるべく近いところからL1以上の大きさの連続空き領域を検索し、圧縮したデータをライトする。この方法によれば、伸長ダーティデータは、ディスクヘライトしようとする度に、その論理アドレスに平均圧縮率を乗じた、アドレスの位置を目標として、その近傍にデータをライトし、リターンする。

【0067】図21により、まとめ書きの準備処理を説明する。ステップ2101で、ディスクヘライトしようとする伸長ダーティデータのあるキャッシュセグメントを、伸長ダーティキューからはずし、キャッシュセグメントを空きとする。それからステップ2102で、圧縮したディスクライト対象の圧縮ダーティデータを、ステップ184で決定した物理アドレスに基づき、物理ディスクシリンドラグループの番号を算出し、その番号の物理ディスクシリンドラグループの圧縮ダーティキャッシュセグメントの物理アドレス順のキューに、データの物理アドレスに基づいてつなぎ、リターンする。

【0068】まとめ書き用の、物理ディスクシリンドラグループ毎の、圧縮ダーティキューに接続された、圧縮ダーティセグメントは、まとめ書き処理（ステップ191）により、リード/ライト処理、及びダーティデータのディスクライト準備処理、とは非同期にディスクへまとめ書きする。このまとめ書きのステップでは、基点となる圧縮ダーティデータ、或いは、基点となる物理アドレスを、カレントなヘッド位置を用いてSSTF (Shortest Seek Time Fast) アルゴリズムや、SCANアルゴリズムを使用して決める、或いは、圧縮ダーティデータの、圧縮時刻順や、データへのホストコンピュータからの最終アクセス時刻が古い順、あるいは管理するキュー中でのデータの並ぶ順序により決めることができる。また、基点の近傍の圧縮ダーティデータの検索は、図8のキャッシュキュー管理テーブルの、物理ディスクセグメントグループの圧縮ダーティキャッシュセグメントの物理アドレス順のキューのポインタから、キューを検索することによって可能である。ここで、圧縮ダーティデータを、物理アドレスで管理するデータ構造としては、物理ディスクに対し、物理シリンドラ毎に該当データの有無と、該当データが有る場合にアロケートするトラック管理テーブルへのポインタとを有するシリンドラ管理テーブルと、シリンドラ毎に、シリンドラ内のトラック上の該当データの有無と、該当データが有る場合、そのデータの管理情報へのポインタを有するトラック管理テーブルと

の、二段のテーブル構造を採用することもできる。

【0069】圧縮機能付きディスクにおいては、ライト時に、圧縮データの長さが変動するので、時間が経つと、論理アドレスと物理アドレスの対応の順序性が失われ、空き領域のフラグメント化が生じる。そのため、いわゆるコンパクションと呼ばれる、空き領域の収集処理や、論理アドレス順に、ディスク上の格納位置を並べ直すソート処理等の、ディスクの再構成処理が、必要である。図22に、ディスク再構成処理の、起動フローを示す。

【0070】本実施例の、実現上の代替案を、図24、図25、図26、図27、図28を用いて、以下に示す。

【0071】図24及び図25は、圧縮ダーティデータを、ディスクにライトする際に、圧縮データのディスク上後方に余分な領域ができた場合、その領域の扱い方を示す。

【0072】図24は、圧縮ダーティデータを、旧圧縮データ上に上書きしたとき（ステップ23）に、新圧縮データの後部にL0-L1の大きさの余分な領域ができた場合に、その領域の全部/または一部を空き領域として開放し（ステップ27）、その空き領域に直接、別のデータを書き込んだり、コンパクション処理によりその空き領域を移動してデータ格納に利用することを可能とする方法を示す。

【0073】図25では、圧縮ダーティデータを、旧圧縮データ上に上書きしたとき（ステップ23）に、新圧縮データの後部にL0-L1の大きさの余分な領域ができた場合に、その領域の全部/または一部を空き領域としないでおき、将来、対応するデータの圧縮データ長が増加した場合に、ディスク上の格納位置の変更をしないで元の領域にデータをライトすることを可能として、データの論理アドレスと物理アドレスの順序性をできるだけ損なわないようにする方法を示す。

【0074】図26では、圧縮ダーティデータのデータ長L1が、旧圧縮データのデータ長L0よりも長い場合（ステップ24）に、旧圧縮データの後方のデータを別領域に移動し（ステップ29）、旧圧縮データの後方に十分な大きさの空き領域を作って旧圧縮データの格納領域と後続の空き領域をつなげて、そこに新圧縮データを書く方法を示している。

【0075】図27は、圧縮ダーティデータのディスクライト処理の延長上で、ディスク上の空き領域集め（コンパクション）処理を起動する方法を示している。図27中で、旧圧縮データ上に、そのままでは新たな圧縮データをライトできない場合に、コンパクション処理を起動し（ステップ2B）、それから旧圧縮データの近傍に必要な大きさの空き領域を見つけてデータをライトする（ステップ2C）。コンパクション処理の起動方法としては、元の領域にそのままではデータがライトできない

事象の発生回数をカウントして閾値比較をする、あるいは、前記事象発生時に、ディスク上のフラグメンテーション度合いをチェックしたうえでコンパクション処理を起動するなど、いくつかの方法が考えられる。

【0076】図28は、圧縮データを、分割してディスクに格納する方法を示している。図28では、そのままでは、新たな圧縮データを、元の領域にライトできない場合に、圧縮データを分割して、ディスク上で分散させて、元の領域の近傍に格納する(ステップ2D)。この方法を用いれば、これまで述べた方法に比べて、ディスクのコンパクション処理を減らすことができる。

【0077】以上により、ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、ホストコンピュータからみえる論理的なデータのアドレスと圧縮されたデータのディスク上の格納アドレスとの間の順序性をできるだけ保つことにより、シーケンシャルリード性能や先読み効果の低下を防止することができるようになる。また、同時にキャッシュ上のディスク未反映のダーティデータのまとめ書きを行うことが可能になる。

【0078】これにより、ビットコストが良い、高性能なディスク装置を実現できる。

【0079】(2)次に、本発明を実施した、圧縮機能とライトバックキャッシュを有する、ディスク制御装置の第二の実施例を述べる。本実施例では、古くなったダーティデータを、ディスクライトに備えて圧縮し、ディスク上の格納位置(物理アドレス)を決めた後、特に圧縮したダーティデータを、物理アドレスから検索可能なキューにつなぐことはしない。まとめ書きデータの検索は、論理アドレス上近傍の圧縮ダーティデータを、検索し、そのデータの物理アドレスをチェックし、物理アドレス上近傍のものだけを選択することにより行う。この方法では、第一の実施例に比べて、プログラム構造が簡単になることと、メモリ使用量を少なくできることが利点である。

【0080】本実施例の大部分は、第一の実施例と同一であるので、違う部分だけを、図29、図30、図31、図32を用いて説明する。

【0081】図31中の、3100は、物理ディスク管理テーブルである。3101は、各物理ディスクの管理単位に、論理データが割当てられているか否かを示すデータ有無フラグであり、3102は、割当てられている論理ディスクセグメントの番号である。図32は、圧縮ダーティキャッシュセグメントのキューの先頭及び末尾のポインタを示している。

【0082】図29は、キャッシュ上の伸長ダーティデータのディスクライト準備の処理を示している。ステップ184までは、第一の実施例の、図18に示す処理と同一である。ステップ184で、圧縮ダーティデータの物理アドレスを決定する。次に、ステップ189へ進

み、当該圧縮ダーティデータに対応するテーブルをセットする。具体的には、図5に示される、論理ディスク管理テーブル7中で、当該データに対応する論理ディスクセグメントエントリの、物理ディスク上の位置情報、即ち、hd\_\_addr75、dlen76、d\_\_fwd\_ptr77、d\_\_bwd\_ptr78をセットする。この実施例では、圧縮ダーティデータは、圧縮ダーティデータ専用の1本のキューにて、管理する。即ち、該当データを、伸長ダーティデータのキューからはずし、圧縮ダーティデータのキューにつなぐ。対応して、キャッシュ管理テーブル8中のc\_\_fwd\_ptr、c\_\_bwd\_ptr等も更新する。

【0083】図30は、本実施例におけるまとめ書き処理を示す。まとめ書き処理では、まずまとめ書き対象の圧縮ダーティデータ或いはまとめ書きの基点とする物理アドレスを決定し(ステップ192)、ステップ192で決定した基点の、論理アドレスを元に、論理ディスク管理テーブル7から、その論理アドレス近傍の論理ディスクセグメントのエントリのcsn72からキャッシュ管理テーブルを引き、更に対応するキャッシュセグメントエントリの、csd83、csst82から、圧縮ダーティデータの有無を調べて、圧縮ダーティデータの有る論理ディスクセグメントを選出し(ステップ193)、更に、論理ディスク管理テーブル7の論理ディスクセグメントエントリのhd\_\_addr75から、基点とする物理アドレスの近傍にあるか否かをチェックして、基点物理アドレス上近傍の、圧縮ダーティデータを、選び(ステップ194)、まとめ書きを行う(ステップ195)。

【0084】まとめ書き対象の圧縮ダーティデータから、論理アドレスを求めることは、キャッシュ管理テーブル8の、ldsn8Aを用いることによって可能である。

【0085】基点の物理アドレスから、対応する論理アドレスを求めることは、図31の物理ディスク管理テーブルを用いて、物理アドレスから物理ディスクに割当てられているデータの論理アドレスを求めることによって可能である。

【0086】また、本実施例では、圧縮したダーティデータは、第一の実施例のように物理ディスクシリンダグループ毎に管理するのではなく、図32に示す、圧縮した時刻順の、圧縮ダーティキャッシュセグメントのFIFOキューによって管理する。

【0087】本実施例により、プログラム構造を簡単にして、メモリ使用量を少なくして、圧縮ダーティデータのディスクへのまとめ書きを行うことができる。

【0088】

【発明の効果】本発明により、ライトバックキャッシュとデータ圧縮機能を採用したディスク制御装置において、ホストコンピュータからみえる論理的なデータのA

ドレスと圧縮されたデータのディスク上の格納アドレスとの間の順序性をできるだけ保つことにより、シーケンシャルリード性能や先読み効果の低下を防止することができる。また、同時にキャッシュ上のディスク未反映のダーティデータのまとめ書きを行うことが可能になる。

【0089】これにより、ビットコストが良い、高性能なディスク装置を実現できる。

【図面の簡単な説明】

【図1】本発明を適用するディスク制御装置の位置付けを示すブロック図。

【図2】圧縮機能付きディスク制御装置のハードウェアのブロック図。

【図3】ディスク上の空きセクタを管理する管理テーブルの説明図。

【図4】ライトバックスケジュール処理のフローチャート。

【図5】制御に使用する論理ディスク管理テーブルの説明図。

【図6】制御に使用するキャッシュ管理テーブルの説明図。

【図7】キャッシュメモリのデータ格納部の説明図。

【図8】キャッシュセグメントのキューを管理するテーブルの説明図。

【図9】ディスク上の空き領域を管理するテーブルの説明図。

【図10】空き領域管理エントリを示す説明図。

【図11】メモリの使用の仕方、即ちメモリ構成を示す説明図。

【図12】ディスクの初期化処理のフローチャート。

【図13】ディスクへのリード／ライトの処理フローチャート。

【図14】ライトミス時の処理フローチャート。

【図15】ライトヒット時の処理フローチャート。

【図16】リードヒット時の処理フローチャート。

【図17】リードミス時の処理フローチャート。

【図18】伸長ダーティデータのディスクライト準備の処理フローチャート。

【図19】まとめ書き処理を示す説明図。

【図20】圧縮ダーティデータの、ディスク上の格納位置、即ち物理アドレスを決定する方法の処理フローチャート。

05 【図21】まとめ書きの準備の処理フローチャート。

【図22】ディスク再構成処理の起動の仕方を示すフローチャート。

【図23】圧縮ダーティデータの物理アドレスの決定方法の代替案の処理フローチャート。

10 【図24】ディスク更新時の後続領域の扱いの一方を示すフローチャート。

【図25】ディスク更新時の後続領域の扱いの一方を示すフローチャート。

15 【図26】圧縮データのディスク格納の一方法を示すフローチャート。

【図27】圧縮データのディスク格納時に、連携して空き領域集め（コンパクション）処理を起動する場合の、処理フローチャート。

20 【図28】圧縮データを分割してディスクに分散格納する一方法の処理フローチャート。

【図29】論理アドレスを用いてまとめ書きデータを見つける場合の、伸長ダーティデータのディスクライト準備の処理を示す説明図。

25 【図30】論理アドレスを用いてまとめ書きデータを見つける場合の、まとめ書き処理の処理フローチャート。

【図31】物理ディスク管理テーブルの説明図。

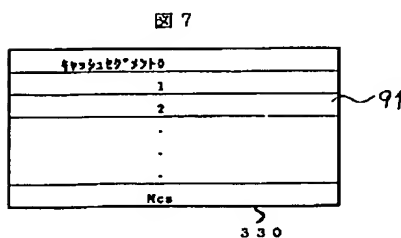
【図32】圧縮ダーティキャッシュセグメントのキューを示す説明図。

【符号の説明】

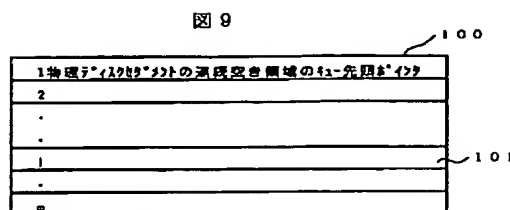
30 181…ダーティデータを選択するステップ、182…判定するステップ、183…選択したデータを圧縮するステップ、184…圧縮したダーティデータのディスク上の格納位置、185…まとめ書き準備を行うステップ、186…対応するデータをディスクから読み出すステップ、187…ディスクから読み出した圧縮状態のデータを伸長するステップ、188…伸長したリードデータを用いて、データを揃えるステップ。

35

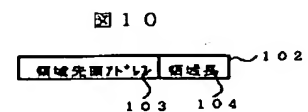
【図7】



【図9】

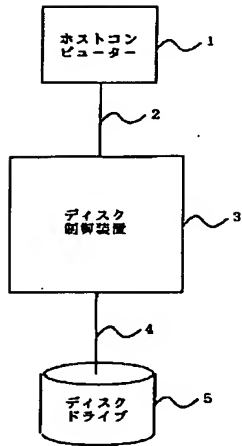


【図10】



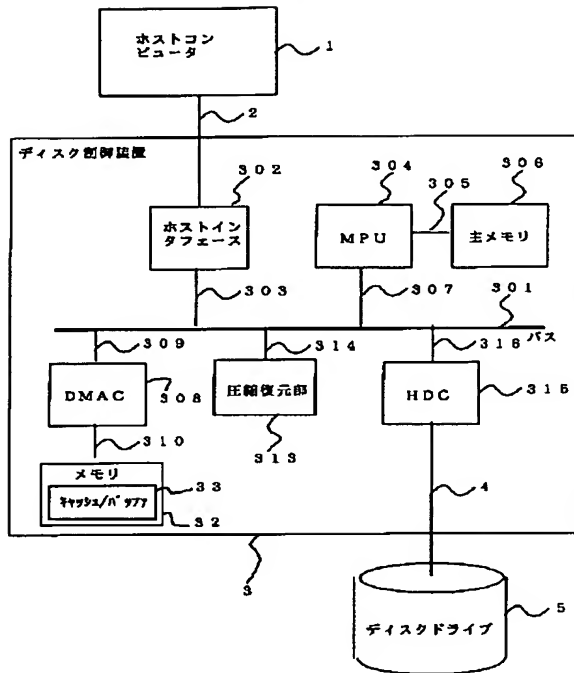
【図1】

図1



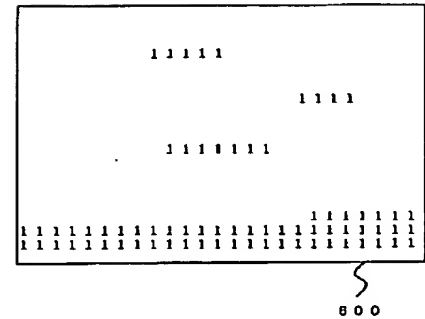
【図2】

図2



【図3】

図3



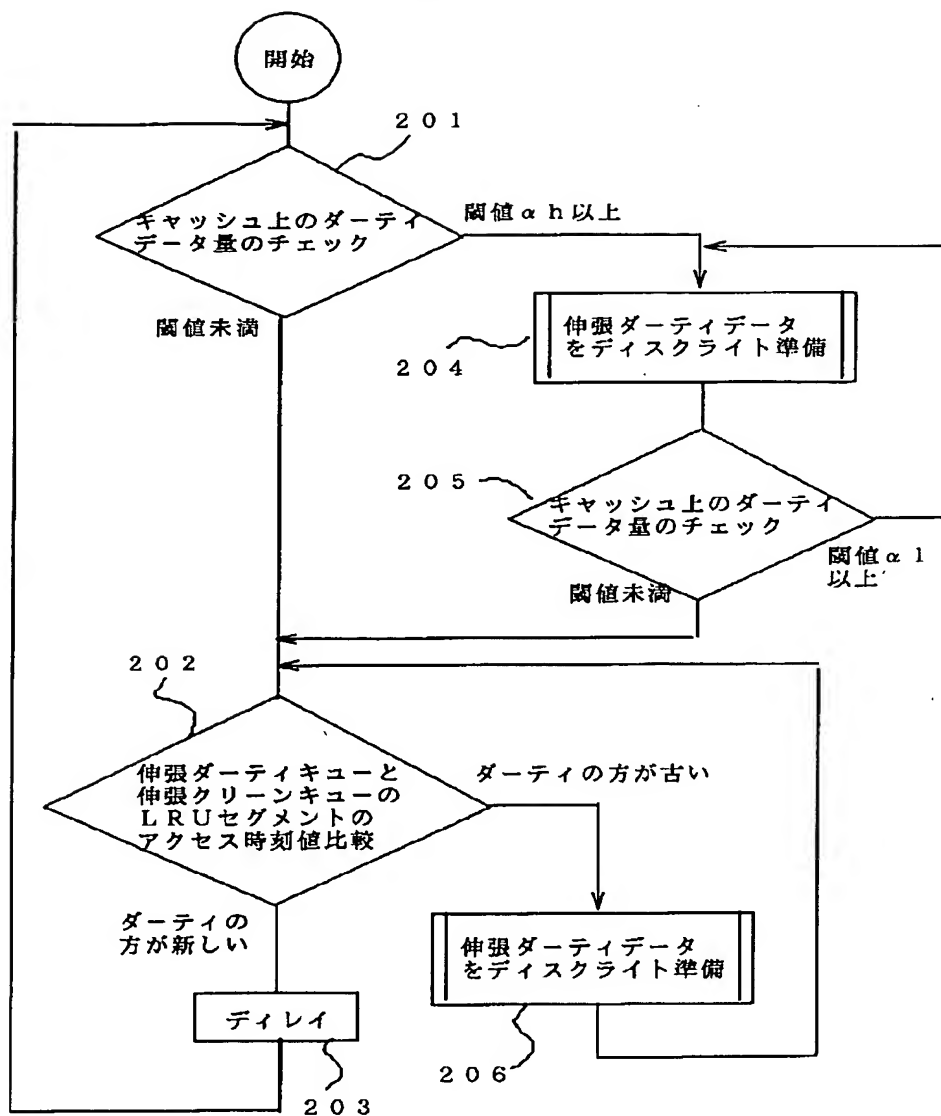
【図8】

図8

空キキャッシュセグメントリスト先頭ポインタ	93
伸張クリーンキャッシュセグメントLRUキュー-MRUポインタ	94
伸張クリーンキャッシュセグメントLRUキュー-LRUポインタ	95
伸張データキャッシュセグメントLRUキュー-MRUポインタ	96
伸張データキャッシュセグメントLRUキュー-LRUポインタ	97
物理ディスクシリンダグループ0の圧縮データキャッシュセグメントの物理アドレス順キュー先頭ポインタ	
1	
2	98
.	
.	
.	
NCG	

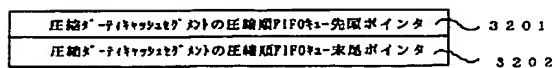
【図4】

図4



【図32】

図32





【図5】

図5

図5 論理ディスク管理テーブル

チャネル管理		物理ディスク管理						
chm(0)	csn(0)	ldu(0)	ldst(0)	hd_addr(0)	dlen(0)	d_fwd_ptr(0)	d_bwd_ptr(0)	
chm(i)	csn(i)	ldu(i)	ldst(i)	hd_addr(i)	dlen(i)	d_fwd_ptr(i)	d_bwd_ptr(i)	
71	72	73	74	75	76	77	78	79

1

.

i

.

Nds

Nds: 論理ディスクの総個数。cha: チャネル判定情報。csn: チャネル番号。

ldu: 当該論理ディスクが使用中か未使用かを示す。

ldst: 当該論理ディスクのデータが物理ディスク上、圧縮されているか否かを示す。

hd\_addr: 当該論理ディスクに対応するコンプレスト論理ディスクの、物理ディスク上の先頭アドレス。

dlen: 当該論理ディスクの物理ディスク上のデータ長。

d\_fwd\_ptr: 前方ポインタ。当該論理ディスクに対応するコンプレスト論理ディスクの、物理ディスク直前のコンプレスト論理ディスクへのポインタ。論理ディスクの番号とする。

d\_bwd\_ptr: 後方ポインタ。当該論理ディスクに対応するコンプレスト論理ディスクの、物理ディスク直後のコンプレスト論理ディスクへのポインタ。論理ディスクの番号とする。

【図6】

図 6

8

キャッシュ・ポインタ	状態情報		伸張時		圧縮時	スタック情報		Idsn
	csu(0)	csst(0)	csd(0)	exist_bits(0)	dirty_bits(0)	access_time(0)	c_fud_ptr(0)	c_bwd_ptr(0)
01ノリ								(0)
1								
.								
1								
.								
Ncs								

csu(i)	csst(i)	csd(i)	exist_bits(i)	dirty_bits(i)	clen(i)	access_time(i)	c_fud_ptr(i)	c_bwd_ptr(i)	Idsn
81	82	83	84	85	86	87	88	89	(1)
80									8A

Ncs:キャッシュ・ポインタの総数。

csu:当該キャッシュ・ポインタが使用中か否かを示す。 csd:当該キャッシュ・ポインタがグ・ディか否かを示す。

csst:当該キャッシュ・ポインタ中のデ・タが、圧縮デ・タか伸張デ・タかを示す。

exist\_bits:キャッシュ・ポインタ中の各デ・タに有効デ・タが存在しているか否かを示す。

dirty\_bits:キャッシュ・ポインタ中の各デ・タにグ・ディデ・タが存在しているか否かを示す。

clen:キャッシュ・ポインタに圧縮デ・タが存在する場合、そのデ・タ長を示す。

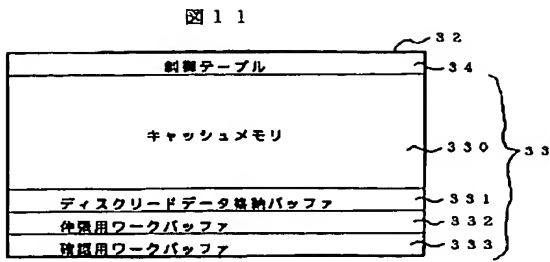
access\_time:キャッシュ・ポインタへの、最新のアクセス時刻を記録する。

c\_fud\_ptr:当該キャッシュ・ポインタが存在するメモリ中の、直前のキャッシュ・ポインタへのポインタ。

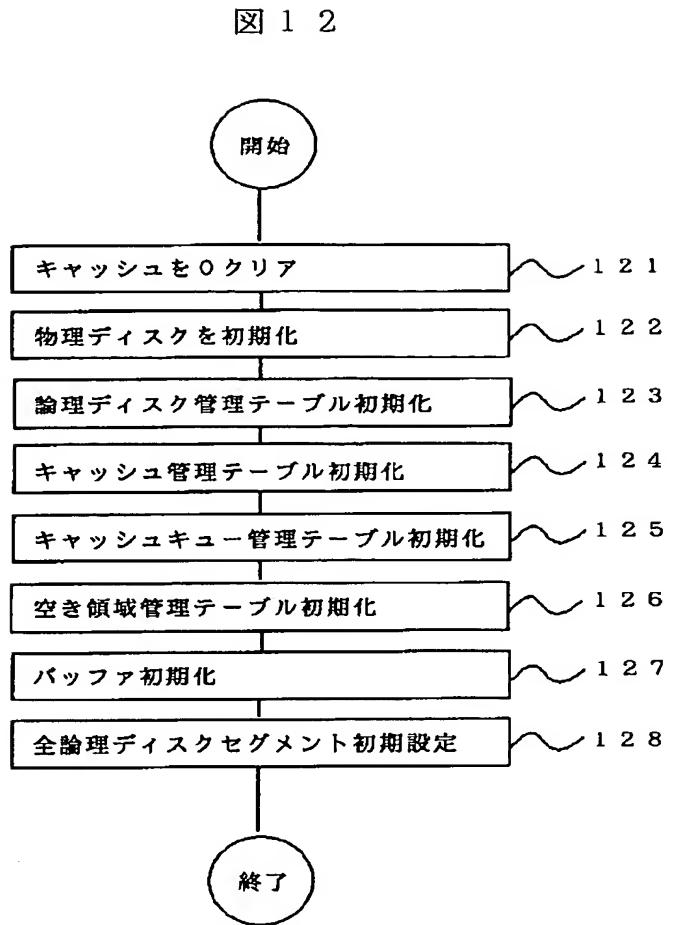
c\_bwd\_ptr:当該キャッシュ・ポインタが存在するメモリ中の、直後のキャッシュ・ポインタへのポインタ。

Idsn:当該キャッシュ・ポインタに割当てられている管理デ・タの番号。

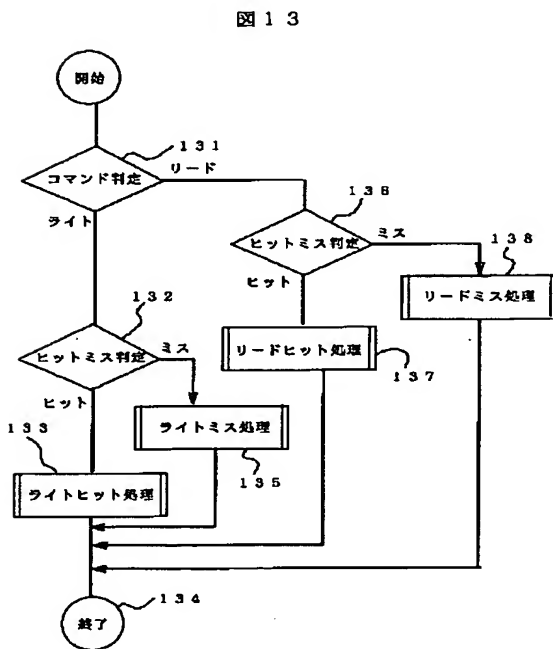
【図11】



【図12】

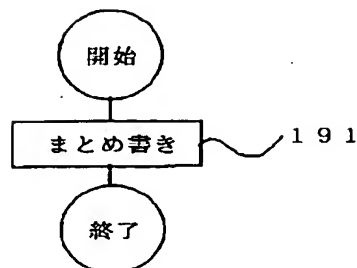


【図13】



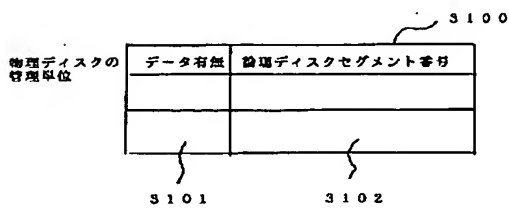
【図19】

図 1 9



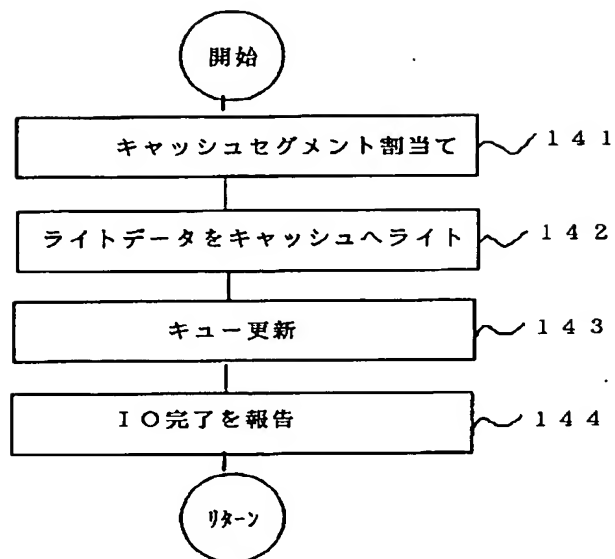
【図31】

図 3 1



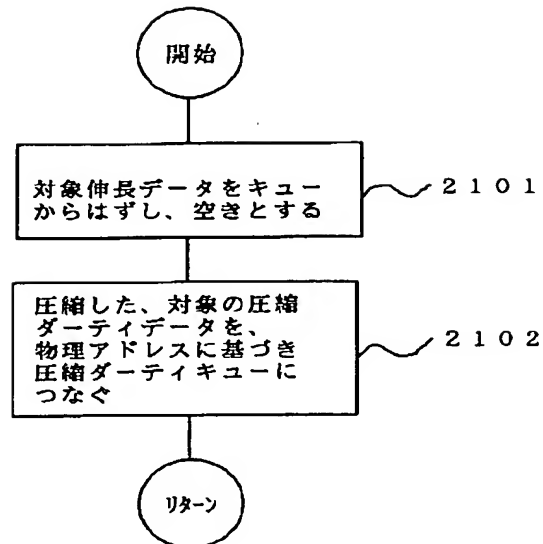
【図14】

図 1 4



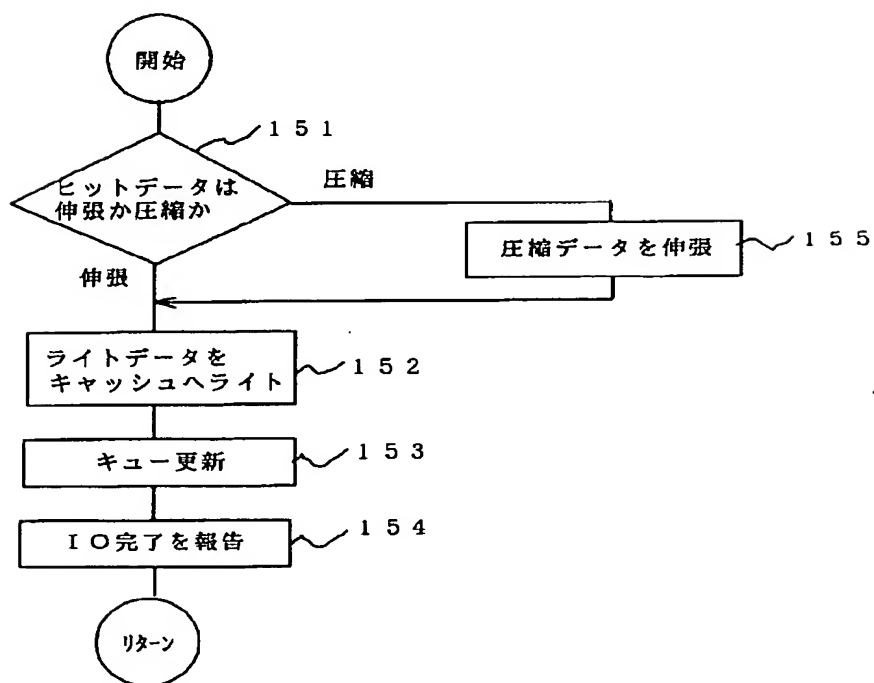
【図21】

図 2 1



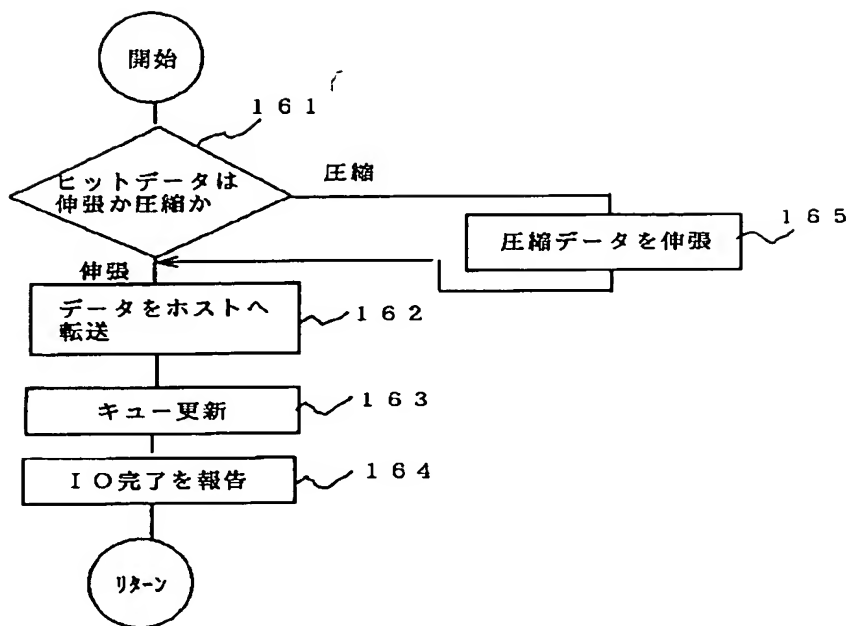
【図15】

図 1 5



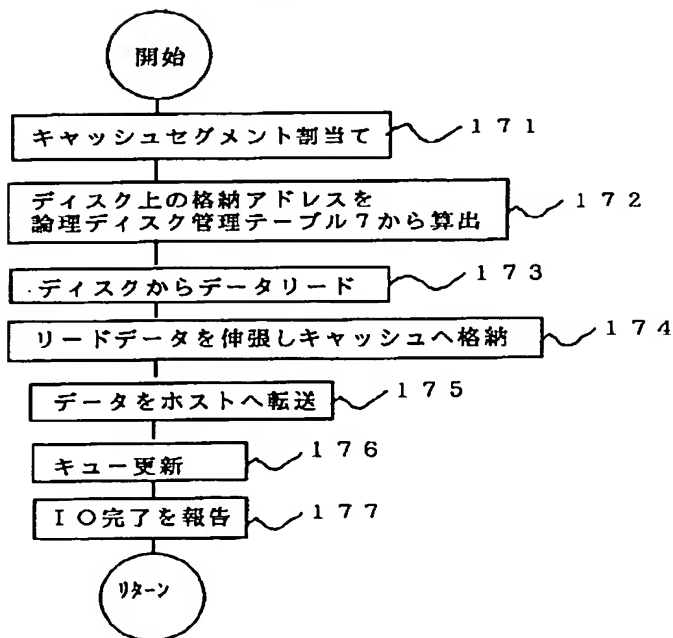
【図16】

図 1 6



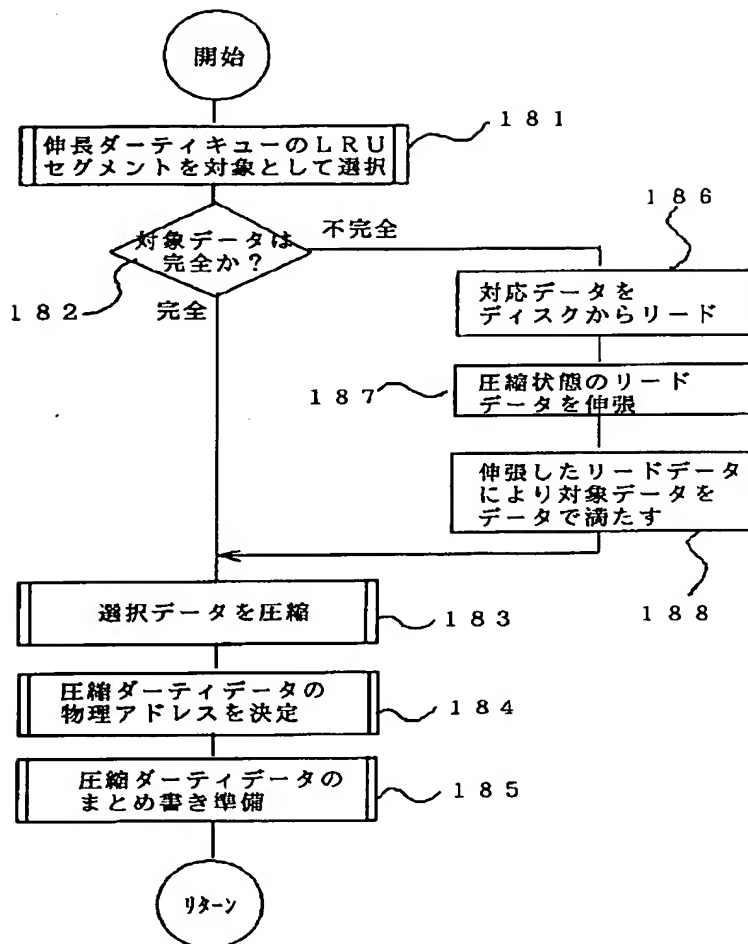
【図17】

図 1 7



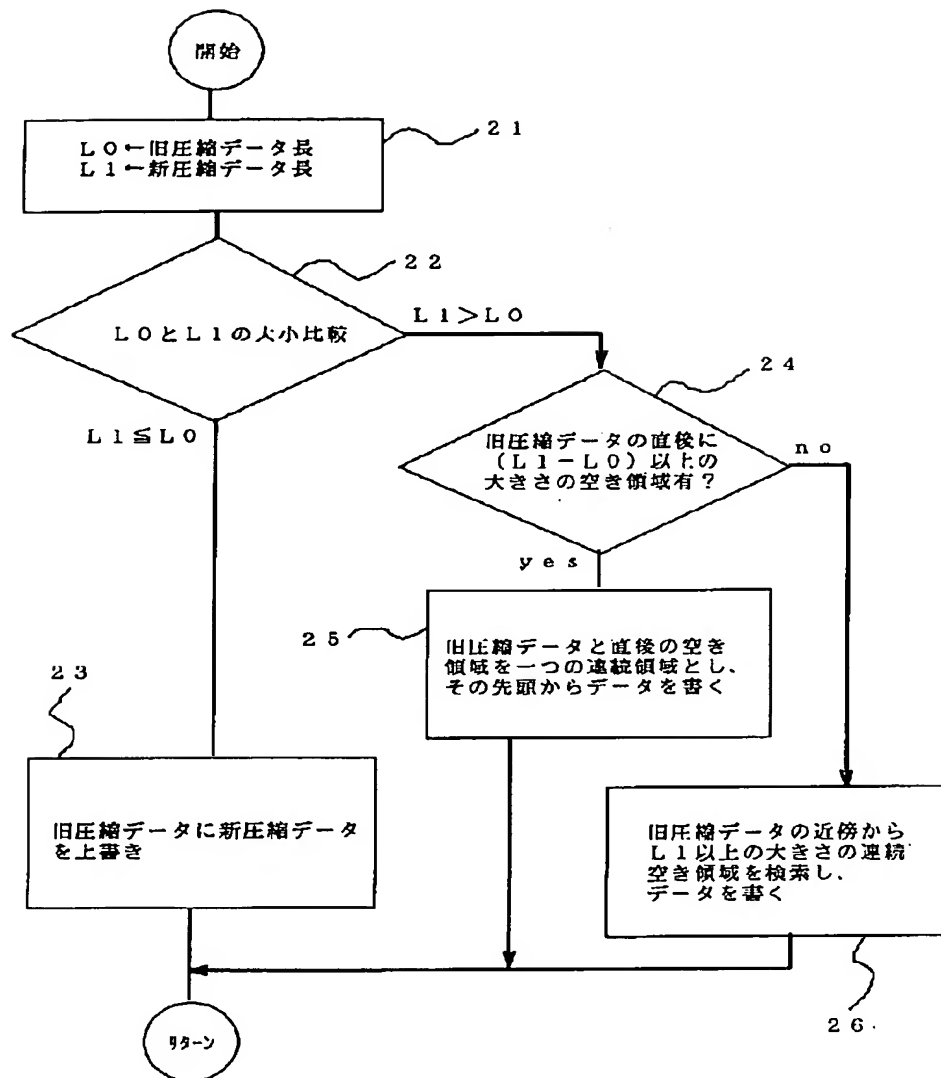
【図18】

図 18



【図 20】

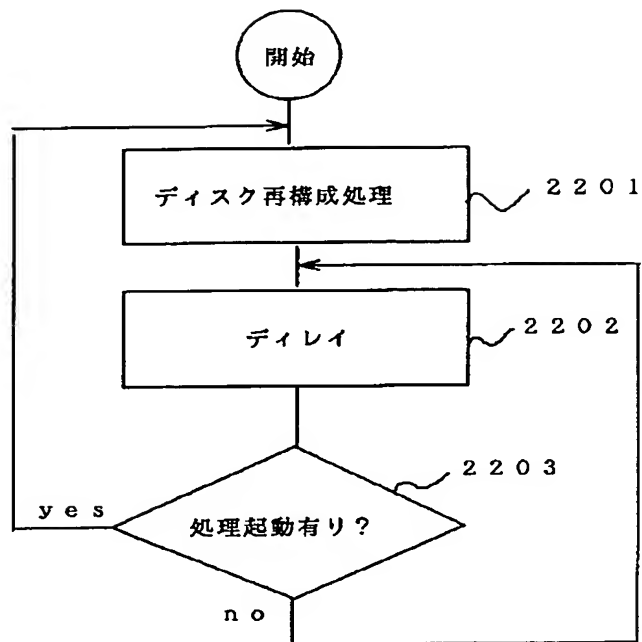
図 20





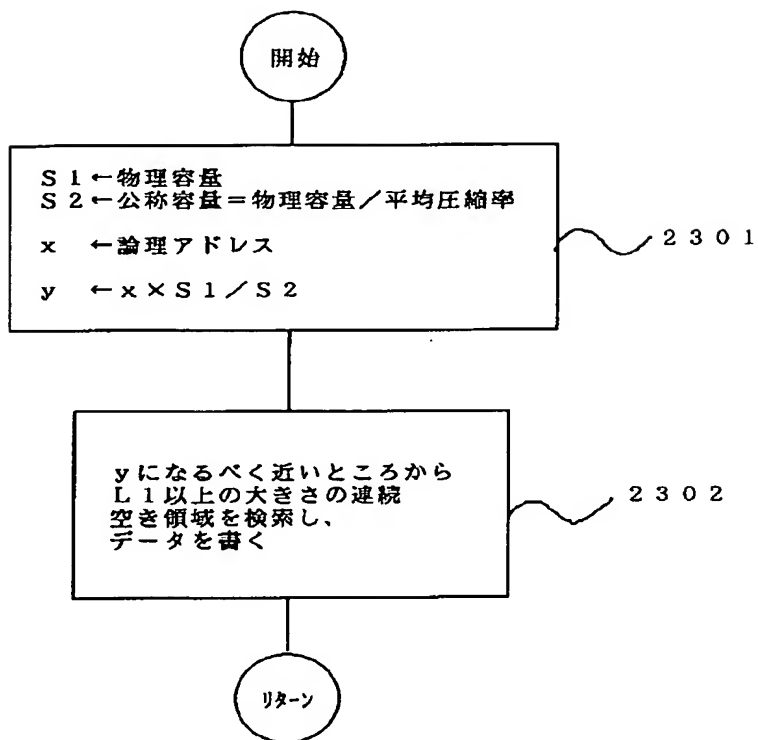
【図 22】

図 22



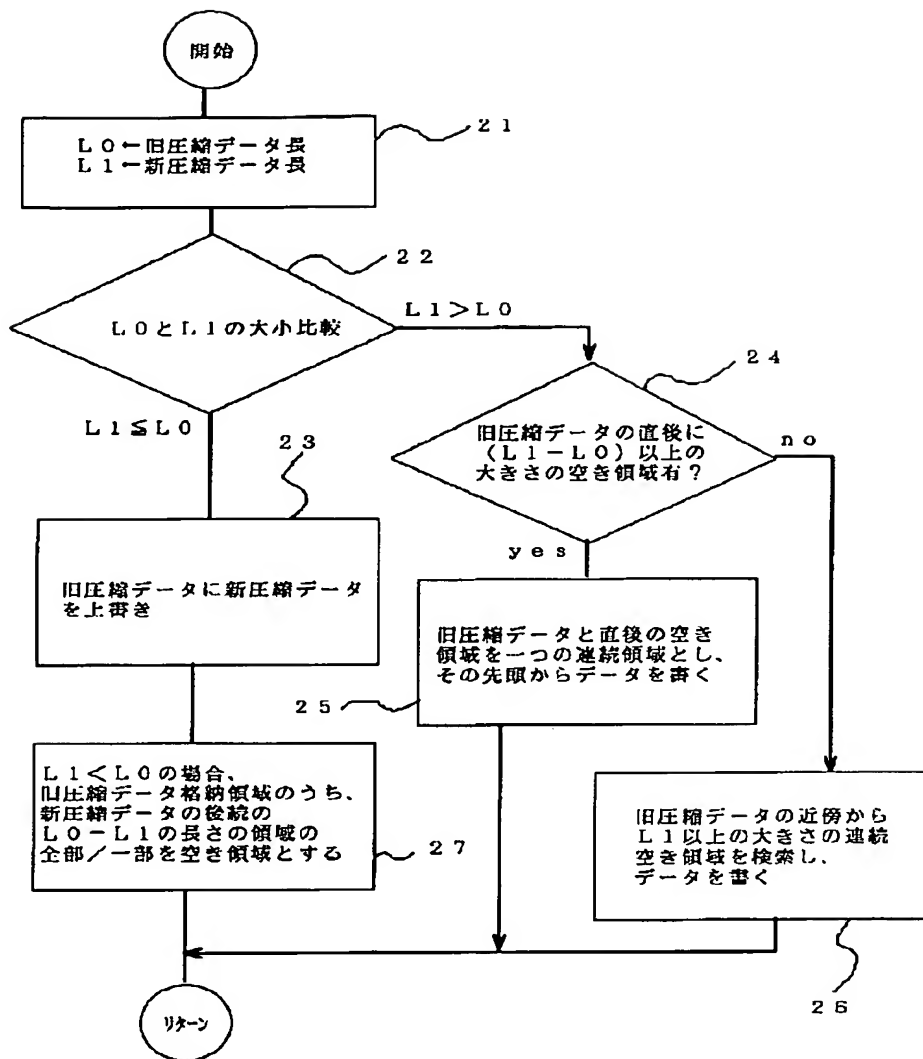
【図23】

図 2 3



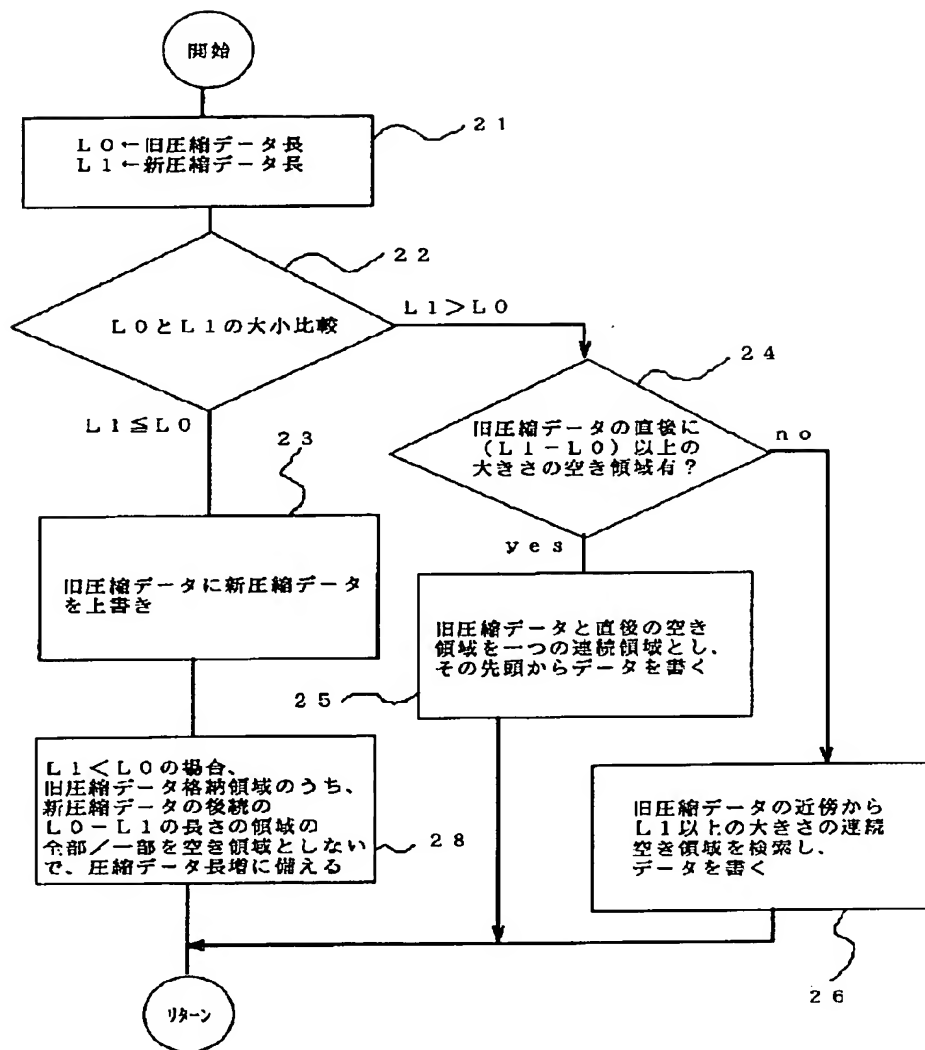
【図24】

図 2 4



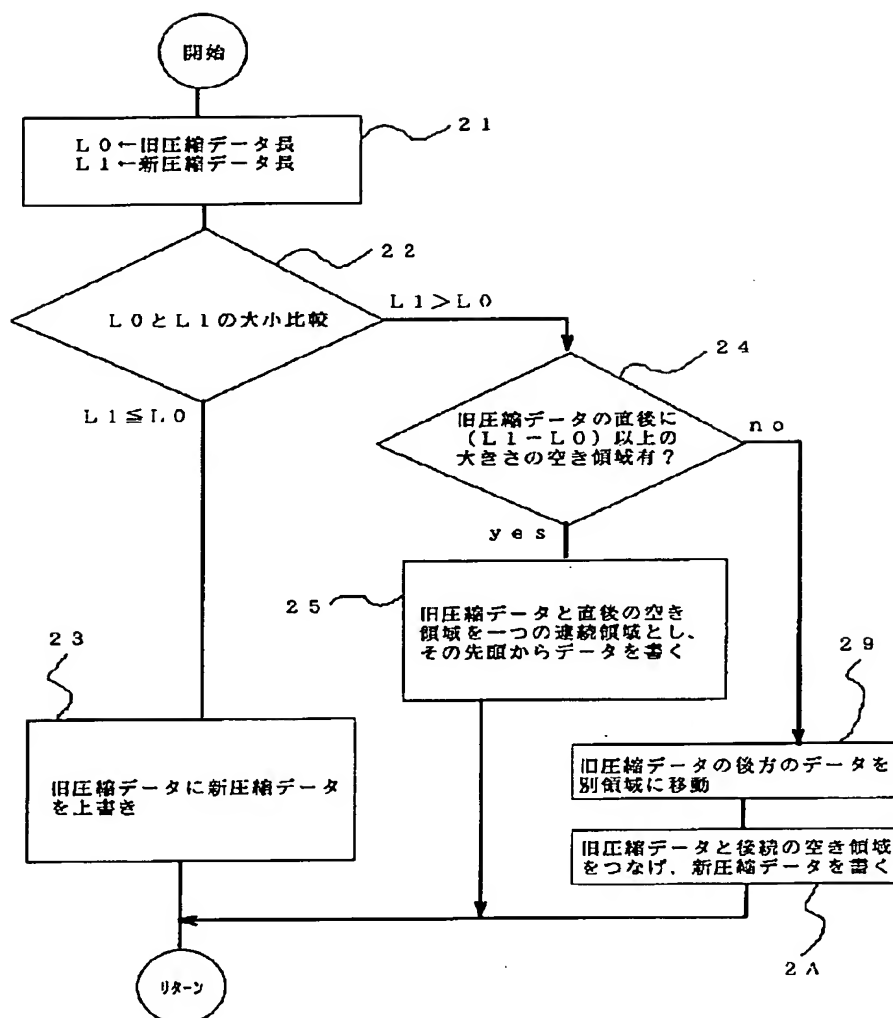
【図25】

図 2 5



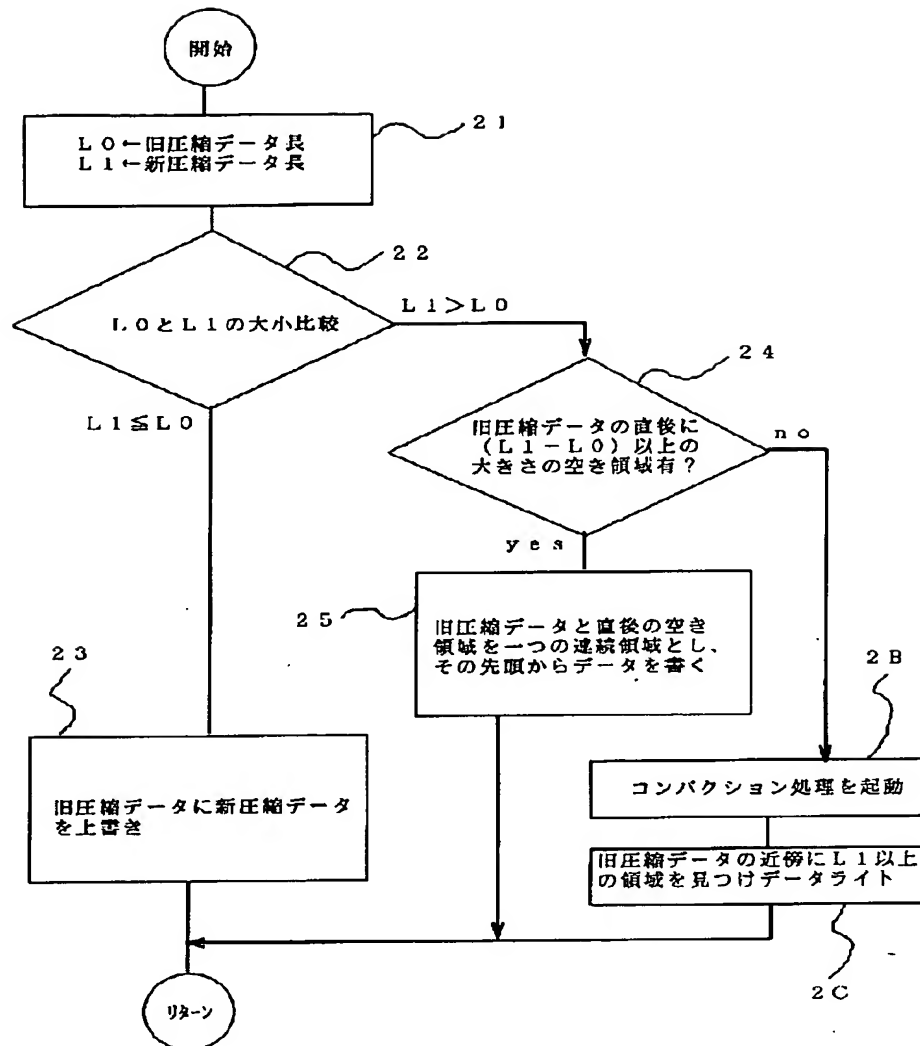
【図 26】

図 26



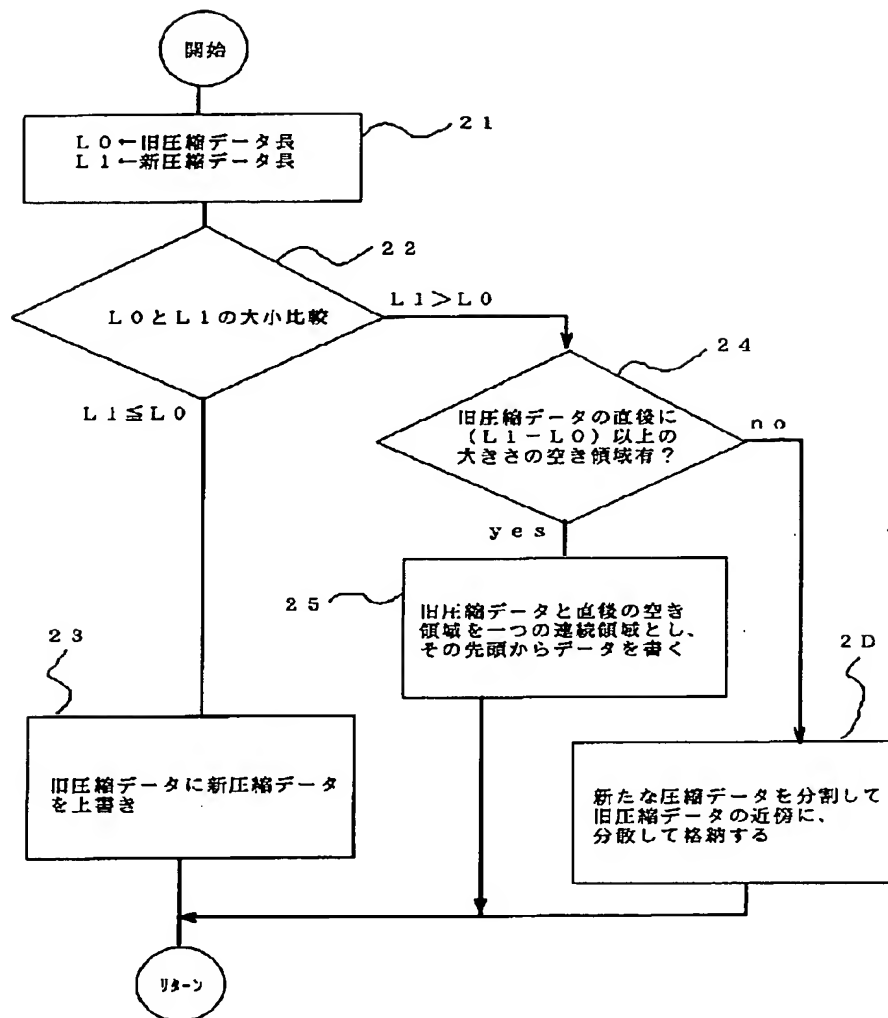
【図 27】

図 27



【図28】

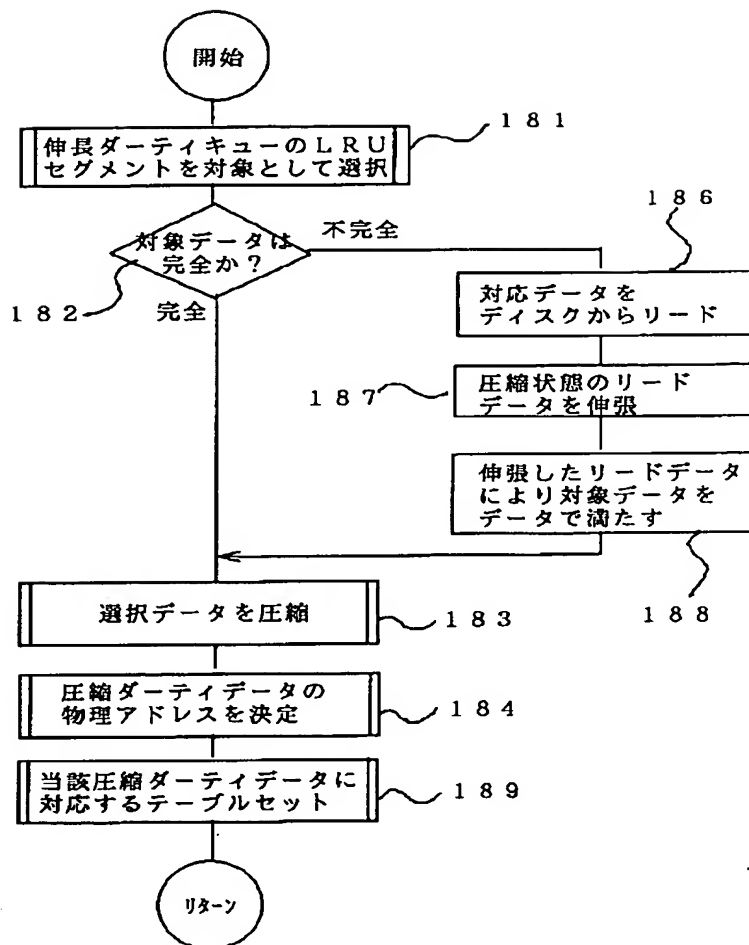
図 2 8





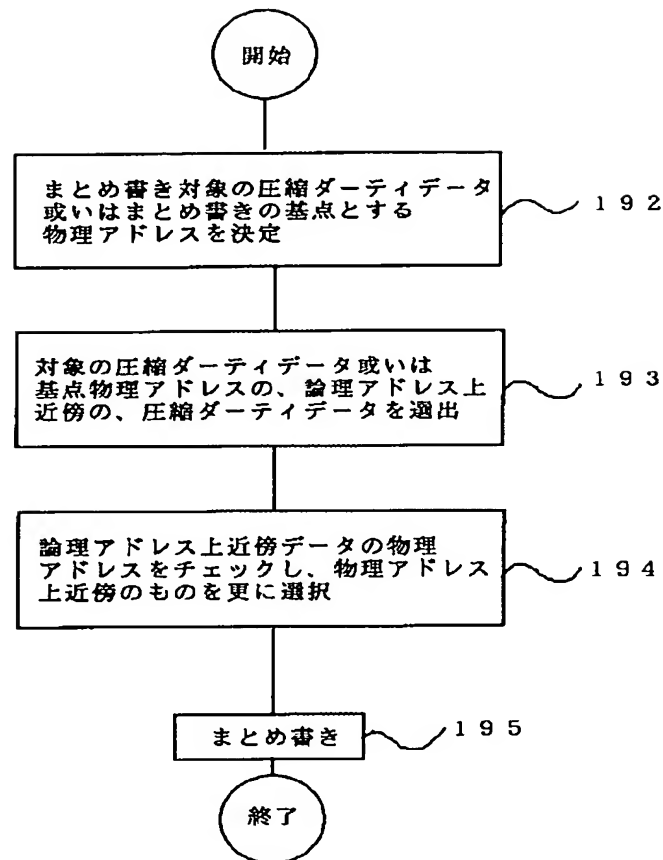
【図 29】

図 29



【図30】

図 3 0



フロントページの続き

(72)発明者 阿知和 恭介  
神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内